

571

| | |
|---|------|
| AATCTCATTCTCCTATGGT TATGGGGCGG GTCATTGATG CCATCACATC GGGGCAATTA | 6540 |
| ACCCAGCAGG ACCTCCTTCT TAGCCTATTT TACTTGCTAC TTGCAGCCTT TGGTATGTAC | 6600 |
| TATTTGCGCT ATGTGTGGCG TATGTATATC CTTGGGACCT CTTATTGCTT GGGACAGATC | 6660 |
| ATGCGGTCTC GCTTGTTTAA GCATTTTACA AAAATGTCGT CAGCCTTTTA TCAAACCTAT | 6720 |
| CGGACGGGTG ATCTGATGGC ACACGCAACC AATGATATCA ATGCCTTGAC TCGTTTAGCA | 6780 |
| GGTGGCGGTG TCATGTCTGC GGTGGATGCC TCTATCACGG CTCTGGTGAC TTTGTTGACC | 6840 |
| ATGCTCTTTA GCATCTCATG GCAGATGACT CTGTTGCCA TTCTCCCCCT ACCTTTTCATG | 6900 |
| GCCTATACGA CTAGTCGCCT AGGGAGAAAG ACTCATAAGG CCTTTGGCGA ATCCCAAGCT | 6960 |
| GCTTTTCTG AACTCAATAA CAAGGTACAG GAGTCCGTAT CAGGTATCAA AGTGACCAAG | 7020 |
| TCTTTCGGTT ATCAGGCAGA CGAGTTGAAG TCTTTTCAGG CAGTCAATGA ATTAACCTTC | 7080 |
| CAAAAGAACC TGCAAACCAT GAAATATGAT AGTCTCTTTG ACCCTATGGT TCTCTTGTTT | 7140 |
| GTTGGTTCGT CCTATGTTTT AACGCTTTTG GTTGGCTCCT TGATGGTTCA GGAAGGGCAG | 7200 |
| ATTACAGTTG GGAATCTAGT CACCTTTATC AGCTATTTGG ATATGCTGGT CTGGCCTCTT | 7260 |
| CTGGCCATCG GTTTCCTCTT TAATACTACT CAGCGAGGGA AGGTTTCTTA CCAGCGGATT | 7320 |
| GAAAATCTTT TGTCTCAGGA ATCTCCTGTA CAAGACCCTG AGTTTCCTCT GGATGGTATT | 7380 |
| GAAAATGGGC GTTTGGAGTA TGCCATTGAC AGCTTTGCTT TTGAAAATGA GGAAACACTG | 7440 |
| ACGGATATTC ACTTTAGTTT GGCAAAAGGG CAAACACTGG GCTTGGTTGG GCAGACAGGC | 7500 |
| TCTGGGAAAA CGTCCTTAAT CAAGCTCCTC TTGCGTGAAT ACGATGTGGA TAAGGGTGCC | 7560 |
| ATTTATCTAA ACGGTCACGA TATTCGGGAC TATCGTCTGA CAGACCTTCG CAGTCTCATG | 7620 |
| GGCTATGTTT CTCAGGACCA GTTCTTTTTT GCGACTTCAA TCCTAGACAA TATCCGCTTT | 7680 |
| GGCAATCCTA ACTTGCCCTT TTCAGCGGTC GAGGAAGCTA CTAAGCTAGC CCGGGTTTAC | 7740 |
| CAAGATATTG TAGACATGCC TCAAGGATTT GATACGCTGA TTGGTGAAAA AGGAGTCACT | 7800 |
| CTTCTGGTG GTCAAAAGCA ACGGTTGGCT ATGAGTCGGG CTATGATTTT AGACCCTGAT | 7860 |
| ATCTTGATTT TGGATGATTC CTTATCCGCC GTAGATGCCA AGACAGAGTA TGCGATTATC | 7920 |
| GACAACCTCA AGGAGATGCG AAAGGACAAG ACAACCATTA TCACTGCCCC TCGCCTCAGT | 7980 |
| GCTGTTGTCC ATGCAGATTT TATTTTAGTT CTACAAAATG GTCAAATTAT CGAACGAGGC | 8040 |
| ACGCACGAAG ACTTGCTAGC TTTGGATGGC TGGTATGCCC AAACCTACCA GTCTCAGCAG | 8100 |
| TTGGAAATGA AAGGAGAAGA AGATGCAGAA TAAACAAGAA CAATGGACTG TATTGAAGCG | 8160 |
| CTTGATGTCT TATCTCAAGC CTTATGGACT CCTGACCTTT TTGGCACTCA GTTTTCTCCT | 8220 |

572

| | | | | | | |
|-------------|-------------|------------|-------------|-------------|-------------|-------|
| AGCGACGACG | GTCATTAAAA | GTGTCATACC | CCTCGTGGCT | TCCCACTTTA | TCGACCAGTA | 8280 |
| TCTCAGCAAT | CTTAACCAAC | TAGCCGTTAC | CGTTTTGCTG | GTCTACTATG | GTCTCTACAT | 8340 |
| CCTACAAACT | GTAGTTCAGT | ATGTCGGCAA | TCTTCTCTTT | GCGCGCGTGT | CTTACAGTAT | 8400 |
| TGTTAGGGAT | ATTCTGTCGGG | ATGCCTTTGC | CAATATGGAG | AAACTGGGCA | TGTCTTACTT | 8460 |
| TGACAAGACG | CCAGCAGGTT | CTATCGTTTC | TCGTTTGACC | AACGATACCG | AGACGATTAG | 8520 |
| TGATATGTTT | TCTGGGATTT | TATCCAGCTT | TATCTCAGCA | GTTTTTATCT | TTCTGACAAAC | 8580 |
| CCTTTATACC | ATGTTGGTGC | TGGATTTTCG | TTTGACGGCT | TTAGTCTTGC | TCTTTCTTCC | 8640 |
| TTTGATTTTC | CTTTTGGTCA | ATCTCTATCG | AAAAAAGTCA | GTGAAATCA | TCGAGAAAAC | 8700 |
| CAGAAGTCTC | TTGTCAGATA | TCAATAGTAA | GCTGGCAGAG | AATATCGAGG | GAATCAGGAT | 8760 |
| TATTTCAGGCC | TTTAATCAAG | AGAAGCGCCT | GCAGGCAGAA | TTTGATGAAA | TCAACCAAGA | 8820 |
| ACACTTGGTC | TACGCCAACC | GTTCTGTAGC | CTTGATGCC | CTCTTTTGA | GACCTGCCAT | 8880 |
| GAGTTTGCTG | AAACTTCTAG | GCTATGCAGT | CTTGATGGCC | TACTTTGGCT | ACCGTGGTTT | 8940 |
| TTCTATCGGG | ATAACGGTCG | GGACCATGTA | TGCCTTTATC | CAGTACATCA | ACCGCCTTTT | 9000 |
| TGACCCCTTG | ATTGAGGTGA | CGCAAACTT | TTCAACTCTG | CAAACGGCTA | TGGTTTCTGC | 9060 |
| AGGTCGTGTC | TTTGCCCTGA | TAGACGAGAG | GACCTATGAA | CCTCTTCAAG | AAAATGGGCA | 9120 |
| AGCCAAAGTC | CAAGAAGGCA | ATATCCGTTT | TGAACATGTG | TGTTTCTCAT | ATGACGGTAA | 9180 |
| ACATCCGATT | CTGGATGACA | TTTCTTTCTC | TGTTAATAAG | GGTGAAACCA | TTGCCTTTGT | 9240 |
| AGGTCATACA | GGTTCAGGGA | AATCGTCTAT | TATCAATGTC | CTCATGCGCT | TTTATGAATT | 9300 |
| CCAGTCAGGG | AGAGTTCTCT | TGGATGATGT | GGATATCAGG | GATTTTCAGTC | AAGAAGAGCT | 9360 |
| GAGAAAAAAC | ATCGGTTTGG | TCTTGCAGGA | ACCCTTCTCT | TATCATGGAA | CTATTAAGTC | 9420 |
| CAATATCGCC | ATGTACCAAG | AAACCACTGA | TGAGCAGGTT | CAGGCTGCGG | CAGCCTTTGT | 9480 |
| GGATGCAGAT | TCCTTTATTC | AAGAACTTCC | TCAGGGGTAC | GACTCCCTCT | TTCCGAGCG | 9540 |
| TGGTTCGAGC | TTCTCTACTG | GGCAACGCCA | GCTTCTTGCC | TTTGCTAGAA | CAGTCGCCAG | 9600 |
| CCAGCCTAAA | ATCCTGATTT | TGGATGAAGC | GACAGCCAAT | ATTGACTCTG | AAACAGAAAG | 9660 |
| CTTGGTTCAA | GCTTCTCTGG | CGAAGATGAG | ACAGGGCCGA | ACAACTATTG | CTATCGCTCA | 9720 |
| CCGCCTTTCT | ACTATTCAAG | ATGCCAACTG | CATCTATGTC | TTGGATAAGG | GACGCATTAT | 9780 |
| CGAGAGTGGA | ACCCATGAGG | AACTCTTGGC | TCTGGGAGGA | ACCTATCACA | AGATGTATAG | 9840 |
| TTTGCAGGCA | GGGGCCATGG | CCGATACTCT | TTGAAAAATCT | CTTTAAACCA | TGTCAGCTTT | 9900 |
| ATCTGCAATC | TCAAAGCTGT | ACTTTGATTT | TCATTGAGTA | CTAGAAGGAA | ATCCTTCAAA | 9960 |
| TTACAGATTT | CTTTCACCGC | CTTTTCCATT | TTGTGGTATA | ATGAAAAATG | TTGACAAATA | 10020 |

573

| | |
|--|-------|
| GTATAATAAA AACAAAGGAG AACAGCATGC TGAAATGGGA AGACTTGCCT GTGGAAATGA | 10080 |
| AATCAAGCGA GGTGAGTCT TACTACCAGC TTGTCTCTAA AAGGAAGGGT TCGCTGATTT | 10140 |
| TCAAGCGTTG CTTGGACTGG GTTTTGGCCT TGGTCTTACT GGTTCGTACC TCTCCCATCT | 10200 |
| TTCTCATCTT GAGCATTGG ATCAAGTTGG ATAGCAAAGG GCCAGTGATT TACAAGCAAG | 10260 |
| AGCGTGTGAC CCAGTACAAC CGTCGGTTCA AGATTGGAA GTTTCGTACC ATGGTGACGG | 10320 |
| ATGCGGATAA AAAAGGAAGT CTGGTGACTT CTGCTAACGA TAGCCGCATT ACCAAGGTTG | 10380 |
| GAAATTTTCAT CCGACGTGTC CGTTTGGACG AACTGCCTCA GTTGGTCAAT GTCCTTAAAG | 10440 |
| GTGAGATGTC CTTGTTCGGT ACACGACCTG AAGTGCCACG TTATACAGAG CAGTATAGCC | 10500 |
| CTGAAATGAT GGCAACCTTG CTCTTGCAAG CAGGGATTAC CTCTCCAGCC AGCATCAACT | 10560 |
| ACAAGGATGA GGACACAATT ATCAGTCAAA TGACGGAGAA AGGTCTGTCA GTTGATCAGG | 10620 |
| CCTATGTGGA GCATGTTCTT CCTGAAAAGA TCGCTATAA CCTCGCCTAT CTCCGAGAGT | 10680 |
| TTAGTTTCTT TGGGGACATC AAAATCATGT TTCAAACCGT GTTTGAGGTA CTAAAATAAA | 10740 |
| GTAGTCATAA GAAATGAGT ACAGATAAAA GGAGCAAATC AATGCCAAAT TACAATATTC | 10800 |
| CATTTTCACC GCCTGATATC ACAGAAGCAG AAATTACTGA AGTAGTGGAT ACCCTGCGTT | 10860 |
| CTGGTTGGAT CACAACAGGT CCTAAAACAA AAGAACTGGA GCGCCGCTTG TCTCTTTACA | 10920 |
| CACAGACACC TAAGACTGTT TGTCTCAACT CTGCGACAGC CGCTCTGGAG TTGATTTTAC | 10980 |
| GCGTTTGGGA AGTGGGACCT GGTGATGAAG TCATCGTTCC AGCCATGACC TATACGGCTT | 11040 |
| CATGTAGTGT CATTACGCAC GTGGGAGCAA CCCCTGTCAT GGTGGATATC CAAGCAGATA | 11100 |
| CGTTTGAGAT GGACTATGAC CTGCTTGAGC AAGCTATCAC TGAGAAAAC T AAGGTGATTA | 11160 |
| TTCCAGTAGA GCTCGCAGGG ATTGTTTGCG ATTATGACCG TTTGTTCCAA GTCGTGGAGA | 11220 |
| AAAAACGTGA CTTCTTTACC GCTTCAAGCA AGTGGCAAAA GGCCTTTAAC CGTATTGTCA | 11280 |
| TTGTCTCTGA TAGTGCCAC GCTTTGGGAT CTATTTATAA AGGACAACCT TCTGGTTCTA | 11340 |
| TCGTGACTT TACTTCCTTC TCATTCCATG CAGTTAAGAA CTTTACAACG GCAGAAGGTG | 11400 |
| GAAGTGCAC TTGGAAGCC AATCCAGTGA TTGATGACGA AGAGATGTAC AAGGAATTCC | 11460 |
| AAATCCTTTC CCTTCACGGG CAACTAAGG ATGCTCTTGC CAAGATGCAA CTGGGGTCAT | 11520 |
| GGGAATACGA TATCGTTACA CCAGCCTATA AGTGCAACAT GACCGATATC ATGGCTTCAC | 11580 |
| TTGGTTTGGT ACAATTGGAC CGCTATCCAA GTTTGTTGCA ACGCCGTAAG GACATTGTGG | 11640 |
| ACCGCTATGA TAGTGGTTTT GCAGGTTCTC GCATCCATCC TTTGGCACAC AAGACTGAAA | 11700 |
| CTGTCAATC TTCACGCCAC CTCTACATCA CCCGTGTAGA AGGAGCAAGC CTAGAAGAAC | 11760 |

574

| | |
|---|-------|
| GCAACCTCAT CATCCAAGAA TTGGCTAAAG CAGGAATTGC AAGTAATGTT CACTACAAAC | 11820 |
| CGCTTCCTCT CTTGACAGCC TATAAGAATC TTGGATTGA TATGACGAAC TATCCTAAGG | 11880 |
| CCTATGCCTT CTTTGAGAAT GAAATTACCC TCCCTCTTCA TACTAAATTA AGCGATGAAG | 11940 |
| AAGTAGACTA TATCATTGAG ACTTTCAAAA CAGTTTCTGA AAAAGTGCTA ACTTTATCAA | 12000 |
| AAAAATGACA AACTACAGTC AAGCGAAAGT GATCCTGCCC CTAAAAAGTC TAATTGAGTG | 12060 |
| TAAAACTGT TGTTTTCAAT TGATAATAGT TTACACCTGT AGTTGAGGCC CCTTCTCCT | 12120 |
| CAGAGAGAGA ATTTTATAG GATTTTCCTT TCTTGTGGGA GTCCCGTGGT TTGAAATAAG | 12180 |
| ATGTGAGCAA TTTAGTGTAG CATTTAGAAT CCTTACTAGA CATCATTTAG AAAATCTAGT | 12240 |
| GTCTTGTCT AGTTTTCAAT TCACCCTATT TTTTGAAAGA CGTGAGTTTC CATGAGTGAG | 12300 |
| ATTGTGGAAT CTCGCGTCTT TTTTGTGTTT CAGAAATATG TTCAAAATTT TGTGCCTGTC | 12360 |
| TTTCATGTTT TAGTCATTCT TTTGCATGAT AGAATTTATA GCATGTTGAT ATTATAATAA | 12420 |
| TACAAATATT CTATATGTTT AGTGATGCTT GCTATACATT ATTAGATCTC CTGCGAGACA | 12480 |
| ATCTATAAAA CACTTGCTTA CGATTACCTA TATGCCCTAT TCCAGTATTT TAGAAGCACT | 12540 |
| GCATCTATTT TTATCGAGGT TAAATCTAGC TTTTATAGAA GGTCTATTTA AGAAATATAT | 12600 |
| TGTAGTGTTT TAGTTTCAAT CCGCCATATG AGCGATATTC AGGTAAATAT CCCTGGCGAA | 12660 |
| TGCTTGATG ACAAGGTATT TGTCTTTTCA TTTATAATTT ACAACATATC AACAAATTTA | 12720 |
| AATATAGTAA ATGGGATATT TTATATTCAA GCTAAGAAAG ATAGCATCAC TTTTGAATGG | 12780 |
| AAGGCTAAAG AGCAAACTAG GAAGTTGGCC ATAGATAGCT CAAAACCCTG CTTTGAGGTT | 12840 |
| GTAGATATAG TAAATGAAA TGAGAATAGG ACAAATTGAT CGGGACAGTC AAATCGATTT | 12900 |
| CTAACAATGT TTTAGAAGTA GAGGTGTACT ATTTTAGTTT CAGTCTACTA TAGAACTGAC | 12960 |
| CAAGTCAGTA ACCTAGACTT AGGCAAGGC GGCCTGACC TAGTTTGAAG AGATTTCCTGA | 13020 |
| AGAGTATAAA TTTTAATATT TTCTTGTTT ATTCCTTGAC AATTCAATTT GGAAATATA | 13080 |
| TGATAAAGAT AATGACAGCG GTGTCATTCT ATCTATTTTA AGAAAAGTAA TAATCAATTG | 13140 |
| TTAAAAATAG TAAAAAATT GGAGGTTCTG ATGAAATATT TTGTTCCG | 13188 |

(2) INFORMATION FOR SEQ ID NO: 71:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 32768 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

575

| | |
|--|------|
| AACGAGTGCA TCAGTCTCAG CAAGCACCAG TGCCTCGGCC TCAGCAAGCA CCAGCGCGTC | 60 |
| TGAATCCGCA TCAACCAGTG CCTCAGCTTC AGCAAGTACC TCAGCATCTG AATCAGCATC | 120 |
| AACAAGTGCA TCGGCTTCAG CAAGCACAAG TGCTTCAGCC TCAGCAAGTA TCTCAGCGTC | 180 |
| TGAATCCGCA TCAACCAGTG CGTCCGCTTC AGCAAGTACT AGCGCCTCAG CATCAGCGTC | 240 |
| AACAAGTGCT TCGGCTTCAG CGTCAACGAG TGCCTCTGAG TCAGCATCAA CGAGTACGTC | 300 |
| AGCCTCAGCA AGCACATCAG CTTCTGAATC TGCATCAACC AGTGCCTCAG CCTCAGCATC | 360 |
| GACAAGCGCC TCAGCTTCAG CAAGTACCAG TGCCTCAGCC TCAGCAAGTA CCAGTGCTTC | 420 |
| AGCCTCAGCG TCGACAAGTG CGTCGGCCTC AACCAAGTGCA TCTGAATCGG CATCAACCAG | 480 |
| TGCGTCAGCC TCAGCAAGTA CTAGCGCCTC AGCCTCAGCA TCAACGAGTG CGTCCGCTTC | 540 |
| AGCAAGTACT AGTGCATCAG CATCAGCATC AACCAAGTGCA TCGGCTTCAG CAAGTACCAG | 600 |
| CGCCTCAGCT TCAGCAAGCA CCAGTGCGTC AGCCTCAGCA AGTACCAGCG CCTCAGCCTC | 660 |
| AGCAAGCACC AGTGCCTCAG CTTCAAGCAAG TACCAGTGGC TCAGCCTCAG CGTCGACAAG | 720 |
| TGCGTCGGCT TCAGCAAGTA CCTCAGCGTC TGAATCAGCA TCAACGAGTG CATCAGCTTC | 780 |
| AGCATCAACA AGTGCTTCAG CTTCAAGCAAG TATCTCAGCG TCTGAATCGG CATCAACGAG | 840 |
| TGCGTCCGCT TCAGCAAGTA CTAGCGCCTC AGCATCAGCG TCAACAAGTG CTTGGGCTTC | 900 |
| AGCGTCAACG AGTGCGTCTG AGTCAGCATC AACCAAGTACG TCAGCCTCAG CAAGCACATC | 960 |
| AGCTTCTGAA TCTGCATCAA CCAGTGCGTC AGCCTCAGCA TCGACAAGCG CCTCAGCTTC | 1020 |
| AGCAAGTACC AGTGCGTCAG CCTCAGCAAG TACCAGTGCT TCAGCCTCAG CGTCGACAAG | 1080 |
| TGCGTCGGCC TCAACCAGTG CATCTGAATC GGCATCAACC AGTGCGTCAG CCTCAGCAAG | 1140 |
| TACTAGCGCC TCAGCCTCAG CATCAACGAG TGCGTCCGCT TCAGCAAGTA CTAGTGCATC | 1200 |
| AGCATCAGCA TCAACGAGTG CATCGGCTTC AGCAAGTACC AGCGCCTCAG CTTCAAGCAAG | 1260 |
| CACCAGTGGC TCAGnCTCAG CAAGTACCAG CGCCTCAGCC TCAGCAAGCA CCAGTGCCCTC | 1320 |
| AGCTTCAGCA AGTACCAGTG CGTCAGCCTC AGCGTCGACA AGTGCGTCGG CTTCAAGCAAG | 1380 |
| TACCTCAGCG TCTGAATCAG CATCAACGAG TGCATCAGCT TCAGCATCAA CAAGTGCTTC | 1440 |
| AGCTTCAGCA AGTACCAGTG CGTCGGCTTC AGCATCAACG AGTGCTTCAG TCTCAGCGTC | 1500 |
| AACCAGTGCC TCTGAATCAG CATCAACAAG TGCCTCGGCT TCAGCAAGCA CCAGTGCGTC | 1560 |
| GGCTTCAGCA AGTACTAGTG CATCGGCTTC AGCATCGACA AGTGCGTCTG AATCGGCATC | 1620 |
| AACGAGTGCT TCGGCTTCAG CATCAACGAG TGCCTCAGCC TCAGCAAGCA CATCAGCTTC | 1680 |
| TGAATCTGCA TCAACCAGTG CGTCCGCTTC AGCGTCAACC AGTGCGTCGG CTTCAAGCGTC | 1740 |

576

| | |
|---|------|
| GACAAGTGCT TCGGCTTCAG CATCAACGAG TCGGTCGGCC TCAGCAAGCG CAAGTACCTC | 1800 |
| AGCGTCAGct TCCGCCTCAA CCAGTGCCTC GGCTTCAGCA AGCACAAGTG CGTCAGCCTC | 1860 |
| AGCAAGTATC TCAGCGTCTG AATCGGCATC AACGAGTGCG TCTGAGTCAG CATCAACGAG | 1920 |
| TACGTACGCC TCAGCAAGCA CATCAGCTTC TGAATCTGCA TCAACCAGTG CGTCAGCCTC | 1980 |
| AGCATCGACA AGCGCCTCAG CTTAGCAAG TACCAGTGCT TCAGCCTCAG CGTCGACAAG | 2040 |
| TGCGTCGGCC TCAACCAGTG CATCTGAATC GGCATCAACC AGTGCCTCAG CCTCAGCAAG | 2100 |
| TACTAGTGCA TCAGCTTCAG CATCAACGAG TGCATCGGCT TCAGCATCAA CCAGTGCCTC | 2160 |
| GGCTTCAGCG TCAACCAGTG CGTCAGCTTC AGCAAGTACC AGTGCTTCAG TCTCAGCATC | 2220 |
| AACAAGTGCT TCAGCCTCAG CATCGACAAG TGCCTCGGCT TCAGCAAGCA CATCAGCATC | 2280 |
| TGAATCAGCG TCAACCAGTG CTTGCGCTTC AGCAAGTACC AGTGCTTCAG CTTAGCATC | 2340 |
| AACCAGCGCC TCGGCTCAG CAAGCACCTC AGCTTCTGAA TCGGCCTCAA CCAGCGCCTC | 2400 |
| GGCCTCAGCA AGCACCTCAG CTTCTGAATC GGCCTCAACC AGCGCCTCAG CCTCAGCATC | 2460 |
| AACGAGTGCT TCGGCTTCAG CAAGCACAAG CGCCTCGGGT TCAGCATCAA CGAGTACGTC | 2520 |
| AGCTTCAGCG TCAACCAGTG CTTAGCCTC AGCATCAACA AGTGCCTCAG CCTCAGCAAG | 2580 |
| TATCTCAGCG TCTGAATCGG CATCAACGAG TCGCTCTGAG TCAGCATCAA CGAGTACGTC | 2640 |
| AGCCTCAGCA AGCACCTCAG CTTCTGAATC GGCCTCAACC AGTGCCTCAG CCTCAGCATC | 2700 |
| GACAAGCGCC TCAGCTTCAG CAAGTACCAG TGCTTCAGCC TCAGCGTCAA CAAGTGCCTC | 2760 |
| GGCCTCAACC AGTGATCTG AATCGGCATC AACCAGTGCG TCAGCCTCAG CAAGTACTAG | 2820 |
| TGCATCGGCT TCAGCATCAA CCAGTGCCTC GGCTTCAGCG TCAACCAGTG CGTCAGCTTC | 2880 |
| AGCAAGTACC AGTGCTTCAG TCTCAGCATC AACAAGTGCT TCAGCCTCAG CATCGACAAG | 2940 |
| TGCCTCGGCT TCAGCAAGCA CATCAGCATC TGAATCAGCG TCGACAAGCG CCTCAGCTTC | 3000 |
| AGCAAGTACC AGTGCCTCAG CCTCAGCGTC GACAAGTGCG TCAGCCTCAG CAAGTACTAG | 3060 |
| TGCATCAGCT TCAGCATCAA CGAGTGCATC GGCTTCGGCG TCAACCAGTG CATCAGAGTC | 3120 |
| AGCAAGTACC AGTGCCTCAG CTTCCGCATC AACAAGTGCC TCGGCTTCAG CAAGCACCAG | 3180 |
| TGCGTCGGCT TCAGCAAGTA CTAGCGCCTC AGCCTCAGCC TCAACCAGTG CGTCAGCCTC | 3240 |
| AGCAAGTATC TCAGCGTCTG AATCGGCATC AACGAGTGCG TCCGCTTCAG CAAGTACTAG | 3300 |
| CGCCTCAGCC TCAGCGTCAA CAAGTGCATC GGCTTCAGCG TCAACGAGTG CGTCTGAATC | 3360 |
| GGCATCAACG AGTGCCTCCG CTTAGCAAG TACTAGCGCC TCAGCCTCAG CGTCAACAAG | 3420 |
| TGCATCGGCT TCAGCATCAA CGAGTGCCTC CGCTTCAGCA AGTACTAGCG CCTCAGCCTC | 3480 |
| AGCGTCAACA AGTGATCGG CTTAGCGTC AACGAGTGCG TCTGAGTCAG CATCAACGAG | 3540 |

577

| | |
|--|------|
| TGCGTCAGCC TCAGCAAGCA CATCAGCTTC TGAATCTGCA TCAACCAGTG CGTCAGCCTC | 3600 |
| AGCATCGACA AGCGCCTCAG CTTAGCAAG TACCAGTGCG TCAGCCTCAG CGTCGACAAG | 3660 |
| TGCGTCGGCT TCAGCAAGTA CCAGTGCGTC AGCCTCAGCA AGTACCAGTG CGTCAGCCTC | 3720 |
| AGCGTCGACA AGTGCGTCGG CCTCAACCAG TGCATCTGAA TCGGCATCAA CCAGTGCGTC | 3780 |
| AGCCTCAGCA AGTACTAGTG CATCAGCTTC AGCATCAACG AGTGCATCGG CTTAGCATC | 3840 |
| AACCAGTGCA TCAGAGTCAG CAAGTACCAG TGCGTCAGCT TCCGCATCAA CAAGTGCCCTC | 3900 |
| GGCTTCAGCA AGTACTAGCG CCTCAGCCTC AGCGTCAACA AGTGCTTCAG CTTCCGCGTC | 3960 |
| AACCAGCGCC TCGGCCTCAG CAAGTATCTC AGCGTCTGAA TCGGCATCAA CAAGTGCCCTC | 4020 |
| GGCTTCAGCA TCAACGAGTG CATCAGTCTC AGCAAGCACC AGTGCGTCGG CCTCAGCAAG | 4080 |
| CACCAGCGCG TCTGAATCCG CATCAACCAG TGCCTCAGCT TCAGCAAGTA CCTCAGCATC | 4140 |
| TGAATCAGCA TCAACAAGTG CCTCGGCTTC AGCAAGCACA AGTGCTTCAG CCTCAGCAAG | 4200 |
| TATCTCAGCG TCTGAATCGG CATCAACGAG TCGCTCCGCT TCAGCAAGTA CTAGCGCCTC | 4260 |
| AGCATCAGCG TCAACAAGTG CTTGGGCTTC AGCGTCAACG AGTGCGTCTG AGTCAGCATC | 4320 |
| AACGAGTACG TCAGCCTCAG CAAGCACATC AGCTTCTGAA TCTGCATCAA CCAGTGCGTC | 4380 |
| AGCCTCAGCA TCGACAAGCG CCTCAGCTTC AGCAAGTACC AGTGCGTCAG CCTCAGCAAG | 4440 |
| TACCAGTGCT TCAGCCTCAG CGTCGACAAG TGCGTCGGCC TCAACCAGTG CATCTGAATC | 4500 |
| GGCATCAACC AGTGCGTCAG CCTCAGCAAG TACTAGCGCC TCAGCCTCAG CATCAACGAG | 4560 |
| TGCGTCCGCT TCAGCAAGTA CTAGTGCATC AGCTTCAGCA AGTACTAGCG CCTCAGCCTC | 4620 |
| AGCGTCGACA AGCGCCTCAG CTTAGCAAG TACCAGTGCG TCAGCCTCAG CGTCGACAAG | 4680 |
| TGCGTCGGCT TCAGCAAGTA CCTCAGCGTC TGAATCAGCA TCAACAAGTG CGTCGGCTTC | 4740 |
| AGCATCAACG AGTGCATCAG CTTAGCATC AACAAAGTGCT TCAGCTTCAG CAAGTACCAG | 4800 |
| TGCGTCGGCT TCAGCATCAA CGAGTGCTTC AGTCTCAGCG TCAACCAGTG CCTCTGAATC | 4860 |
| CGCATCAACA AGTGCCCTCG CTTAGCAAG CACCAGTGCT TCGGCTTCAG CGTCAACGAG | 4920 |
| TGCGTCTGAG TCAGCATCAA CGAGTGCGTC AGCCTCAGCA AGCACATCAG CTTCTGAATC | 4980 |
| TGCATCAACC AGTGCGTCAG CTTCCGCATC AACAAAGCGCC TCGGCCTCAG CAAGTACAAG | 5040 |
| TGCTTCAGCC TCAGCATCAA CCAGTGATC AGCTTCAGCC TCAACAAGTG CTTAGCCTC | 5100 |
| AGCGTCAACC AGTGCCCTCG CTTAGCAAG TACCAGTGCG TCAGCTTCAG CAAGCACAAG | 5160 |
| TGCGTCAGCT TCAGCATCAA CCAGTGCTTC GGCTTCGGCA TCAACAAGTG CCTCAGCATC | 5220 |
| AGCATCAACG AGTGCGTCAG CCTCAGCAAG TACTAGTGCA TCAGCATCAG CATCAACCAG | 5280 |

| | | | | | |
|------------|-------------|------------|------------|-------------|-----------------|
| 578 | | | | | |
| TGCATCAGCC | TCAGCAAGTA | TCTCAGCGTC | TGAATCGGCA | TCAACGAGTG | CATCAGCATC 5340 |
| AGCATCAACG | AGTGCAATCGG | CTTCAGCGTC | AACCAGTGCA | TCAGTCTCAG | CAAGCACCAG 5400 |
| TGCGTCGGCT | TCAGCATCAA | CGAGTGCCTC | AGCCTCAGCA | AGTATCTCAG | CGTCTGAATC 5460 |
| GGCATCAACG | AGTGCGTCAG | CCTCAGCAAG | TACTAGTGCA | TCGGCTTCAG | CAAGCACCAG 5520 |
| TGCGTCGGCT | TCAGCATCAA | CCAGTGCCTC | AGCCTCAGCA | AGTATCTCAG | CGTCTGAATC 5580 |
| GGCATCAACG | AGTGCGTCAG | CCTCAGCAAG | TACTAGTGCA | TCAGCmTCAG | CATCAACGAG 5640 |
| TGCATCGGCT | TCAGCAAGTA | CCAGCGCCTC | AGCTTCAGCA | AGCACCAGTG | CGTCAGCCTC 5700 |
| AGCAAGTACC | AGCGCCTCAG | CCTCAGCAAG | CACCAGTGCC | TCAGCTTCAG | CAAGTACCAG 5760 |
| TGCGTCAGCT | CAGCATCAAC | AAGTGCTTCA | GCTTCGGCCT | CAACAAGTGC | GTGAGCTTCA 5820 |
| GCATCAACGA | GTGCGTCGGC | TTCAGCAAGC | ACCAGTGCCT | CGGCCTCAGC | AAGCACCAGT 5880 |
| GCTTCAGCTT | CAGCATCAAC | AAGTGCGTCA | GCTTCAGCAA | GTACATCAGT | TTCAAATTCA 5940 |
| GCAAACCATT | CGAACTCACA | AGTTGAAAT | ACTTCTGGAT | CGACAGGTAA | ATCCCCAAAA 6000 |
| GAATTCGCTA | ATACAGGTAC | TGAGTCGTCA | ATTGGATCTG | TGTTACTTGG | AGTTCTAGCA 6060 |
| GCTGTTACAG | GTATTGGATT | GGTTGCGAAA | CGCCGTAAAC | GTGATGAAGA | AGAGTAAGAC 6120 |
| AACCTGTAAC | GTTAGGCTAA | ACTAACTCGC | GCACATAAAT | CAAGGAGAAA | ATTGCTAGTG 6180 |
| GATGATAAAA | TAACAGTCAT | TGTACCAGTA | TACAATGTGG | AAAACTATCT | GAGGAAGTGC 6240 |
| CTAGATAGTA | TTATTACTCA | AACATATAAA | AATATTGAGA | TTGTTGTCGT | TAATGATGGT 6300 |
| TCTACGGATG | CTTCAGGTGA | AATTTGTAAA | GAATTTTCAG | AAATGGATCA | CCGAATCTC 6360 |
| TATATAGAAC | AAGAAAATGC | TGGTCTTTCT | GCCGCACGAA | ACACCGGTCT | GAATAATATG 6420 |
| TCCGGAATTT | ATGTGACCTT | TGTGGACTCG | GATGATTGGA | TTGAGCAAGA | TTATGTAGAA 6480 |
| ACTCTATATA | AAAAAATAGT | AGAGTATCAG | GCTGATATTG | CAGTTGGTAA | TTATTATTCT 6540 |
| TTCAACGAAA | GTGAAGGAAT | GTTCTACTTT | CATATATTGG | GAGACTCCTA | TTATGAGAAA 6600 |
| GTATATGATA | ATGTTTCTAT | CTTTGAGAAC | TTGTATGAAA | CTCAAGAAAT | GAAGAGTTTT 6660 |
| GCTTTGATAT | CTGCTTGGGG | TAAACTCTAT | AAGGCAAGAT | TGTTTGAGCA | GTTGCGCTTT 6720 |
| GACATAGGTA | AATTAGGAGA | AGATGGTTAC | CTCAATCAAA | AGGTATATTT | ATTATCAGAA 6780 |
| AAGGTAATTT | ATTTAAATAA | AAGTCTTTAT | GCTTATCGGA | TTAGAAAAGG | TAGTTTATCA 6840 |
| AGAGTTTGGA | CAGAAAAGTG | GATGCACGCT | TTAGTTGATG | CTATGCTCTGA | ACGTATTACG 6900 |
| CTACTAGCTA | ATATGGGTTA | TCCTCTAGAG | AAACACTTGG | CAGTTTATCG | TCAGATGTTG 6960 |
| GAAGTCAGTC | TCGCCAACGG | TCAAGCTAGT | GGTTTATCTG | ACACAGCAAC | GTATAAAGAG 7020 |
| TTTGAAATGA | AACAAAGGCT | TTTAAATCAG | CTATCGAGAC | AAGAGGAAAG | TGAAAAGAAA 7080 |

579

GCCATTGTCC TCGCAGCAAA CTATGGCTAT GTAGACCAAG TTTTAACGAC AATCAAGTCT 7140
ATTTGTTATC ATAATCGTTC GATTTCGTTT TATCTGATC ATAGCGATTT TCCAAATGAA 7200
TGGATTAAGC AATTAAATAA GCGCTTAGAG AAGTTTGA CTGAGAAATTAT TAATTGTCGG 7260
GTAACCTCTG AGCAAATTC ATGTTATAAA TCGGATATTA GTTACACAGT CTTTTTACGC 7320
TATTTTCATAG CTGATTTTCGT GCAAGAAGAC AAGGCCCTCT ACTTGGACTG TGATCTAGTT 7380
GTAACGAAAA ATCTGGATGA CTTGTTTGCT ACAGACTTAC AAGATTATCC TTGGGCTGCT 7440
GTTAGAGATT TTGGGGGCAG AGCTTATTTT GGTCAAGAAA TCTTTAATGC CGGTGTTCTC 7500
TTGGTAAACA ATGCTTTTTG GAAAAAGAG AATATGACCC AAAAATTAAT TGATGTAACC 7560
AATGAATGGC ATGATAAGGT GGATCAGGCA GATCAGAGCA TCTTGAATAT GCTTTTTGAA 7620
CATAAATGGT TGGAATTGGA CTTTGATTAT AATCATATTG TCATTCATAA ACAGTTTGCT 7680
GATTATCAAT TGCCTGAGGG TCAGGATTAT CCTGCTATTA TTCACTATCT TTCTCATCGG 7740
AAACCGTGA AAGATTTGGC GGCCCAAACC TATCGTGAAG TTTGGTGGTA CTATCATGGG 7800
CTTGAATGGA CAGAATTGGG ACAAACCAT CATTACATC CATTACAAAG ATCTCACATC 7860
TATCCAATAA AGGAACCTTT CACTTGTCTA ATCTATACTG CCTCAGACCA TATTGAACAA 7920
ATTGAGACAT TGTTTCAATC CTTGCCTGAT ATTCACTTTA AGATAGCAGC TAGAGTAATA 7980
GTTAGTGATC GATTGGCTCA GATGACAATT TATCCAAACG TGACTATATT TAACGGAATT 8040
CACTATTTGG TAGATGTCGA TAATGAATTG GTAGAAACCA GTCAAGTACT TTTAGATATT 8100
AATCATGGCG AAAAGACAGA AGAAATCTC GATCAATTTG CTAATCTTG CAAGCCTATC 8160
TTATCCTTTG AAAATACTAA AACCTATGAA GTAGGTCAGG AGGCATATGC TGTGACCAA 8220
GTTCAAGCAA TGATTGAAAA ATTGAGAGAA ATAAGCAAAT GAAGAAAAAT CATTTAGTAG 8280
GAGATGCTCT GATTTTGACG GTTAGTGATC AGATTGAAGA GTTGGATTAT TTTTATAAA 8340
ATTTCTCCGT TCATCATATA TGAAAGTTGT TCAAACATCA GAGTGCTTTA TAAAAATAA 8400
ATAGACCTAA AGATATTTAA TATGAACTGC ACCCAAAG TTAGACAGAA AAAATCTAAC 8460
TTTTTGGsGT CAGTACAATA TTAGGGTGTG ATTAATTATC TTTTLAGGTG AAAATGATTC 8520
TATATTATAG CTGTTTGATA CGAAATTTAT TATAAGGAAA TTATGTTAAT GAATACAAAA 8580
TCTATAGTTT TTAATGCAGA TAATGATTAT GTAGATAAAT TAGAACTGC AATTAAATCT 8640
ATTTGTTGTT ATAATAATTG TTTAAAATTT TATGTATTTA ATGATGATAT TGCCTCAGAG 8700
TGGTTTTTGA TGATGAATAA GCGATTGAAG ACTATACAA CTGAAATCGT TAATGTAAAG 8760
ATTGTAGATC ATGTTCTTAA AAAGTTTCAT TTACCGTTAA AGAATTTAAG TTATGCCACT 8820

580

| | | | | | | |
|------------|------------|------------|-------------|------------|-------------|-------|
| TTCTTTCGTT | ATTTTATACC | TAATTTTGTC | AAAGAAAGTC | GTGCTTTATA | CCTAGATTCT | 8880 |
| GACATCATTG | TTACAGGAAG | TTTAGACTAT | TTATTTTGATA | TAGAACTAGA | TGGTTATGCC | 8940 |
| TTGGCAGCAG | TAGAAGATTC | TTTTGGTGAT | GTTCCCTTCTA | CCAATTTTAA | CTCCGGAATG | 9000 |
| TTATTAGTTA | ATGTAGATAC | TTGGAGAGAT | GAAGATGCTT | GTTGCAAACT | GTTAGAACTG | 9060 |
| ACCAATCAAT | ATCATGAAAC | AGCATATGGA | GATCAAGGAA | TTTTAAATAT | GTTATTCCAT | 9120 |
| GATAGATGGA | AAAGATTAGA | CCGAAATTTT | AATTTTATGG | TGGGGATGGA | TAGCGTCGCA | 9180 |
| CACATAGAAG | GAAATCATAA | ATGGTATGAG | ATTTCTGAGT | TGAAAAATGG | AGATTTACCT | 9240 |
| AGTGTTATAC | ATTATACTGG | GGTAAACCT | TGGGAAATAA | TTTCCAATAA | TCGCTTTAGA | 9300 |
| GAAGTTTGGT | GGTTTATAA | TCTGTTAGAA | TGGTCTGATA | TTTTATTGAG | AAAAGACATT | 9360 |
| ATTAGTCGTA | GTTTCGAAGA | ACTTGTATAC | AGTCCTAAAG | CTCATACAGC | AATTTTACA | 9420 |
| GCTAGTTGTG | AGATGGAGCA | TGTAGAATAT | TTGATAGAAA | ATTTACCAGA | GGTACATTTT | 9480 |
| TCTATACTAG | CACATACATA | TTTTGCGTCT | AGTGTCGTTG | CTTTATTAAG | ATATAGCAAT | 9540 |
| GTTACGATTT | ATCCTTGTTT | TTCTCCATTT | GATTATCGAA | AAATTTTGGA | TAATTTAGAT | 9600 |
| TTTTATTTAG | ATATTAATCA | TTATAAAGAA | GTGGATAATA | TTGTATCCGT | TGTTCAACAA | 9660 |
| CTATCTAAAC | CAATTTTAC | CTTTGAAAAA | ACTAGTCATG | ATATAGGCAA | TCAAACCTAAT | 9720 |
| ATATTTTCTT | CAACCGAACC | AAACAAAATG | GTAGAGGCTA | TTAGACAATT | TATAGGAGAA | 9780 |
| TAAGTTTATG | GCAGACGAAC | TAATTAGTAT | TGTAGTTCCA | ATCTACAACG | TTGAGAATTA | 9840 |
| TTTGCGAATG | TGTTTGATA | GCATTCAGAA | TCAGACGTAT | CAAAATTTTG | AGTGTTTATT | 9900 |
| AATCAATGAT | GGCTCTCCAG | ATCATTATC | CAAAATATGT | GAAGAATTTG | TAGAGAAAGA | 9960 |
| TTCTCGTTTC | AAATATTTTG | AGAAAGCAAA | CGGCGGTCTT | TCATCAGCTC | GTAACCTAGG | 10020 |
| TATTGAATGT | TCGGGGGGGG | GCGTACATTA | CTTTTGTAAG | CTCTGATGAT | TGGTTGGAAC | 10080 |
| ATGATGCTTT | AGACCGATTA | TATGGTGCTT | TGAAAAAGGA | AAACGCAGAT | ATTAGTATCG | 10140 |
| GGCGTTATAA | TTCTTATGAT | GAAACACGCT | ATGTGTATAT | GACTTATGTT | ACGGATCCAG | 10200 |
| ATGATTCTCT | AGAAGTGATA | GAAGGTAAAG | CAATTATGGA | TAGGGAAGGT | GTCGAAGAAG | 10260 |
| TCAGAAATGG | GAACTGGACT | GTAGCTGTCT | TGAAGTTATT | CAAGAGAGAG | TTACTACAAG | 10320 |
| ATTTACCATT | TCCTATAGGA | AAAATTGCAG | AGGATACTTA | CTGGACATGG | AAGGTACTTC | 10380 |
| TAAGAGCTTC | GAGGATAGTC | TATTTGAATC | GTTGTGTTTA | CTGGTACCGT | GTTGGTTTAT | 10440 |
| CTGATACTTT | ATCGAATACA | TGGAGTGAAA | AGCGTATGTA | TGATGAAATT | GGGGCTAGGG | 10500 |
| AAGAAAAGAT | AGCTATTTTA | GCAAGTTCAG | ACTATGACTT | GACCAATCAT | ATTTTGATTT | 10560 |
| ATAAAAATAG | ATTACAAAGA | GTGATAGCAA | AATTAGAAGA | ACAAAATATG | CAGTTCACAG | 10620 |

581

| | |
|--|-------|
| AGATTACAG AAGAATGATG GAAAAATTGT CTTTACTTCC GTAGATAGTA ATAAAAATG | 10680 |
| AGATAGCGTA ATATGAAACT ACATTTAACA AATTTATACG GCATGGCTGG TGATAGTACG | 10740 |
| GTTATCTTAG CTCAAAATGC TGTTCAAAAG ATAGCTAGTC AACTGGGATT TAGAGAGGTT | 10800 |
| GGTATTTATT TTTACAACAT TGCTTCAGAT AGTCCTTCTG AAATGAATAA GCGTCTGGAT | 10860 |
| GGTATTATGG CCAGTATCTC TATTGGGGAT ATTTTAGTCT TTCAGTCTCC AACCTGGAAT | 10920 |
| GGTTTGAAT TTGATCGTCT CTTGTTTGAT AAGCTAAAGG ATATGCAGGT GAAAATTATT | 10980 |
| TGCTTTATCC ATGATGTTGT TCCCCTCATG TTTGATAGTA ACTATTATCT CATGAAAGAT | 11040 |
| TATCTGTATA TGTATAATCT ATCAGATGTT TTGATAGTGC CGTCAGAGAG AATGAAAACA | 11100 |
| CGCCTGATGG AAGAAGGATT GACGACTAAG AAGATTCTTG TTCAAGGGAT GTGGGATCAT | 11160 |
| CCTCATGATT TATCCTTATA CACCCTGCT TTTAAAAAG AACTTTTTTT TGCTGGAAGT | 11220 |
| TTAGAGCGTT TTCCAGACTT ACAAATTGG TCTCAAGATA CGCCTTTGAG AGTATTTTCA | 11280 |
| AATAAAGGGG AAGCTAGTTC TAGTGCTAGA AGTCTCAGCA TCGAAGGATG GAAAAAGAT | 11340 |
| GAGGAATTGT TGCTAGAATT ATCAAAGGGT GGATTGGGCC TTGTCTGGGG AACCCATCAA | 11400 |
| AATGAGGGAG AAAGTAACCA ATACTATACC TTGAATATAT CTCATAAGGT GAGTACCTAT | 11460 |
| CTAACAGCGG GCATTCCAGT CATGTGACCA AGTAGCTTGT CAACTGCTAA ATTTATAGTA | 11520 |
| GATCAAGGCT TGGGCTTTAT GGCGGATAGT CTGGAAGAGG TTCATGAGAT AGTTGATAAA | 11580 |
| ATGAATCTAC AAGAATATCA AGAAATGACG AATCGTATCA AGACCTTTAG CTATTTGTTA | 11640 |
| AAAGAGGGCT ATTTCACTAA AAAGTTATTG GTAGATGCAA TCTATCACTT GGAATTGAT | 11700 |
| TAAGGGAATG AAATGAACAA AACAATTGTA CTAGCAGGGG ATCGCAATTA CACCAGGCAG | 11760 |
| TTAGAAACAA CGATAAAATC TATTTTATAC CACAATCGAG ATGTTAAGAT TTATATTTTG | 11820 |
| AATCAAGATA TCATGCCAGA TTGGTTTCGC AAACCACGAA AAATAGCTCG CATGTTAGGT | 11880 |
| AGTGAGATTA TCGATGTTAA ACTACCTGAA CAACTGTGT TTCAAGATTG GAAAAGCAA | 11940 |
| GATCACATTA GTAGCATTAC TTATGCTAGA TATTTTATTG CAGATTATAT CCAAGAAGAT | 12000 |
| AAGGTTTTAT ATTTAGACAG TGATTTGATT GTAAATACTT CTTTAGAGAA ATTATTTAGT | 12060 |
| ATTTGTTTAG AAGAAAAATC ACTCGCAGCA GTTAAAGATA CAGATGGAAT TACATTTAAT | 12120 |
| GCAGGTGTTT TATTAATCAA CAATAAAAAA TGGCGTCAAG AGAAATTAAA AGAACGACTA | 12180 |
| ATTGAACAGA GCATTGTTAC AATGAAGGAA GTTGAAGAAG GCCGTTTCGA GCATTTTAAAT | 12240 |
| GGTGATCAAA CGATTTTAA TCAGGTCTTG CAAGATGATT GGTTAGAACT AGGTGAGCT | 12300 |
| TATAATTTAC AAGTAGGGCA TGATATTGTG GCTTTGTATA ACAATTGGCA GGAACATCTG | 12360 |

582

| | |
|--|-------|
| GCTTTTAATG ATAAACCAGT GGTGATTCAT TTTACGACCT ACAGAAAACC CTGGACTACC | 12420 |
| TTGACAGCCA ATCGTTATCG TGATTTATGG TGGGAATTCC ATGATTGGA GTGGAGTCAG | 12480 |
| ATTTTACAAC ACCATATGGG AGAATTTGAA CTAATATCGC CTCTAGATAA GGAATTTTCT | 12540 |
| TGCTTAACCT TAACGAATTC CCAAGATTTA GAAGGAATAG AAGAGCTAGT TACAGCTCTA | 12600 |
| CCTGAGGTGG TATTTTCATAT CGCAGCTTGG ACGGATATGG GAGATAAATT AAAAAAATTA | 12660 |
| GCTGTATATA ATAATGTGAG ATTGCATCCA CAAATTGTTT CACCGGTCTT AGATAAGCTG | 12720 |
| AAAAAGTCAA CAAATCTATA TTTGGATATC AATCATGGTA GTGCAGATGA GAACTTTTTA | 12780 |
| AAATCTTTCG AAGAACAAGA AAAAACGCTA CTAGCTTTTC AATCGACTCA GCACGGAGAG | 12840 |
| TTAGGACAAA TCGTTTTCGA AAATGGGAAA GTTTCCTTTA TGATTGATAC GATTAAAGAT | 12900 |
| TTTAAGAAAA ACGGACATCT TACCTGTTTT CGACAACCTC CAAGTTTAAC TTGTTTAACG | 12960 |
| TTTACGGCTT CTCAGTATAT CGAACAATTG GATTACTTGG CTGGACAGTT GCCAAATGTT | 13020 |
| GTTTTTCAAA TTGCTGCTTG GACAGCTATG GGGCCAAAAT TATATGATTT GTCTAATCGT | 13080 |
| TATCCTAATA TTCAGCTCTA TCCGGCAATT TCTAGAGATA AGCTAGACGA GTTGAAGGAG | 13140 |
| AAGATGGATG CTTATTTAGA TATCAACCTA CTGACTTCAA CATCCGATAT CGTTGCAGAA | 13200 |
| ATGGCTCATC TATCTAAACC TATACTAGCC TTTTATAAAT CTCAAAATGG GAATAATGGC | 13260 |
| CAAAGGTTGT ATTCAAGTGA ACATCCTGAA CGAATGTTGG CTGATTTGCA AAAATTGATA | 13320 |
| ACTAAGGATA TGCTAGAAAA ACCGCTTGAT ATAATCCAGG TGAAAGGGAT AGATGAAACC | 13380 |
| TTGGATTATA TTATTGAACA CAACTCTTCT TTAGTTCGTT TTGGAGATGG GGAAATCAAT | 13440 |
| ATGCTTGCAG GGCATTCAAT TCCCTACCAG GATTATGATG AAGAGTTGGT TTCAATCATG | 13500 |
| AGGGACATTA TCGGCCAAGA AAGTCGAGAA GATTTAGTAG TGTGCCTTCC TGATGCTTTT | 13560 |
| ACAGATCGTT TTAGGTTTAC ATCGTGGGCG ATTCCATTTT GGAAAGATCA CATGGATCAT | 13620 |
| TATATGGATT TTTACAGAGA GTTATGCAGT GATTCATGGT ATGGCTCAAC CTTTGTATCT | 13680 |
| CGCCCTTATA TCGATTTTGA AGACAAGAGT CAAGCTAAAG CTCAATTGTA AAAATTGAAA | 13740 |
| AGCATTTGGG AAAACCGTGA CTTACTGATA GTCGAAGGTG CGACTTCTCG TTCAGGTGTC | 13800 |
| GGAAATGATT TATTCGATGA GGCAAATTCT ATTAAGCGAA TTATCTGTCC TTCTCATAGT | 13860 |
| GCCTTTTCTA GAGTTCATGA ACTTGAACAA GAAATTGAAA AGTATGCTGG TGGTCGCTTG | 13920 |
| ATTTTATGTA TGCTTGACC TACAGCAAAA GTTCTGAGTT ATAATCTATG CCAGATGGGC | 13980 |
| TATCAAGTTT TGGATGTAGG CCATATTGAC TCAGAGTATG AATGGATGAA AATGGGAGCT | 14040 |
| AAAACTAAGG TTAAATTTTC TCATAAACAT ACTGCAGAAC ATAATTCGA CCAAGATATT | 14100 |
| GAATTTATTG ATGATGAAAC CTATAACAGT CAGATTGTTG CACGAATATT AACTAGACT | 14160 |

583

| | |
|--|-------|
| ATTTAAATA AATGATAAGG ATTTAAATG AGAAATACCA AACGCGCTGT AGTATTTGCA | 14220 |
| GGTGATTACG CTTATATTCG ACAAATCGAA ACGGCGATGA AGTCACTCTG TAGACACAAT | 14280 |
| AGTCATTTGA AAATTTATCT GCTAAATCAG GACATTCCTC AGGAATGGTT TAGTCAAATA | 14340 |
| AGAATATATT TACAAGAGAT GGGGGCGAC TTGATTGACT GCAAGTTAAT TGGCTCACAG | 14400 |
| TTTCAAATGA ATTGGTCTAA TAAATTACCT CATATCAATC ATATGACATT TGCACGCTAT | 14460 |
| TTTATTCAG ATTTGTAAAC AGAAGATAAA GTTCTCTATC TAGATAGTGA TTTGATTGTG | 14520 |
| ACTGGTGATT TGACCGATTT GTTTGAATTA GACTTAGGTG AAAATTATTT GGCAGCAGCT | 14580 |
| CGTCTTGCT TTGGAGCAGG AGTCGGCTTC AATGCTGGTG TTCTCTTGAT TAACAACAAA | 14640 |
| AAATGGGGAT CTGAACTAT TCGACAAAAA TTGATTGACT TAACAGAAAA AGAACATGAG | 14700 |
| AATGTGGAAG AAGGAGACCA GTCAATTTTG AATATGTTGT TTAAAGATCA ATATAGTTCC | 14760 |
| CTTGAAGATC AATATAATTT TCAAATAGGA TATGATTATG GGGCGGCAAC CTTTAAACAT | 14820 |
| CAATTCATTT TTGATATTCC GCTCGAACCA CTGCCACTAA TTTTACACTA TATTTCTCAG | 14880 |
| GATAAGCCTT GGAATCAATT TTCTGTTGGA CGTCTAAGAG AAGTTGGTG GGAATACTCT | 14940 |
| TTGATGGATT GGTCTGTTAT TTAAATGAA TGGTTTTCAA AGAGTGTGAA GTACCCTAGT | 15000 |
| AAATCACAAA TATTTAAGTT GCAATGTGTT AATTTAACGA ATTCTTGGTG TGTGAGAAA | 15060 |
| ATCGATTATT TGGCGGAGCA ATTGCCAGAA GTTCATTTTC ATATTGTTGC TTATACAAAT | 15120 |
| ATGGCAAATG AACTACTAGC TTAAACGCGT TTCTCTAATG TTACCGTATA TCCAAATTCC | 15180 |
| TTACCAATGT TATTGGAACA AATAGTAATA GCTTCAGATT TGTATTTGGA TTTGAATCAT | 15240 |
| GATCGAAAAT TAGAAGATGC ATATGAGTTT GTGCTTAAGT AAAAAAACC AATGATAGCT | 15300 |
| TTGACAATA CTTGCTCTGA AAATCTTTCT GAGATTTTCT ATGAAGGTAT CTATCCAAGC | 15360 |
| TCCATTCCGA AAAAAATGGT TGCAGCAATC AGATCTTACA TGAGGTAGAG AACAGTATGA | 15420 |
| GAAAAATCAAT AGTATTAGCG GCAGATAATG CCTATCTTAT TCCTTTAGAG ACGACTATAA | 15480 |
| AGTCTGTATT GTATCACAAT AGAGATGTTG ATTTTATAT TCTCAACAGT GATATAGCTC | 15540 |
| CTGAATGGTT TAAATTATTG GGGAGAAAAA TGGAACTTGT GAATTTCTACA ATTGCGAGTG | 15600 |
| TACACATTGA TAAAGAACTT TTTGAAAGCT ATAAAACAGG ACCTCATATA AATTATGCTT | 15660 |
| CTTACTTTAG ATTTTGTGCG ACAGAAGTGG TTGAATCTGA TAGGGTATTG TATCTGGATT | 15720 |
| CCGATATCAT TGTAAGTGGG GAACTAGCTA CTTTGTGTTGA GATAGATCTC AAAGGATATT | 15780 |
| CAATTGGTGC TGTGATGAT GTCTATGCCT ATGAAGGACG AAAATCTGGA TTTAATACTG | 15840 |
| GTATGTTACT AATGGATGTT GCAAAGTGA AAGAACATTC TATTGTCAAT AGTTTATTTGG | 15900 |

584

| | |
|--|-------|
| AATTAGCGGC CGAGCAGAAT CAAGTTGTTC ATCTTGGGGA TCAGAGTATT TTAAATATTT | 15960 |
| ATTTTGAGGA TAATTGGCTA GCCTTAGATA AAACATATAA TTATATGGTG GGTATTGATA | 16020 |
| TTTATCACCT TGCTCAAGAA TGTGAACGTC TAGATGACAA TCCACCTACA ATTGTTCACCT | 16080 |
| ATGCTAGTCA TGATAAACCT TGGAAACAT ATAGTATATC TAGACTACGT GAATTATGGT | 16140 |
| GGGTTTATAG AGATTGGAT TGGTCAGAGA TTGCTTTTCA ACGTTCCGAT TTAAATTATT | 16200 |
| TTGAAAGAAG CAATCAGTCT AAAAAACAAG TGATGCTTGT GACATGGAGT GCAGATATAA | 16260 |
| AACATTTAGA GTATTTAGTA CAACGGTTAC CTGATTGGCA TTTTCATTTG GCTGCACCGT | 16320 |
| GTGATTGTTC TGAGGAGCTG ACCTCTCTAT CACAGTATAC GAATGTAACA GTATATCAAA | 16380 |
| ATGTATTACA TAGTAGAATT GATTGGCTAT TGGACGATTC TATAGTTTAT TTAGATATTA | 16440 |
| ATACAGGTGG AGAGGTTTTT AATGTAGTTA CAAGGGCACA AGAAAGTGGC AAGAAAATCT | 16500 |
| TCGCTTTTGA TATCACACGT AAAAGTATGG ATGATGGACT CTATGACGGT ATTTTTCTG | 16560 |
| TGGAGAGACC AGATGATTTA GTGGATAGAA TGAAGAATAT AGAGATAGAG TAATGAGTGA | 16620 |
| ATTAATTAGT GTTGTGGTAC CGATATACAA TACGGGAAA TATTTAGTGG AGTGTGTCGA | 16680 |
| GCATATTCTG AAGCAAACCT ATCAAAATAT AGAAATTATT TTAGTTGATG ACGGTCTAC | 16740 |
| GGATAATTCT GGGGAAATTT GTGATGCTTT TATGATGCAA GATAATCGTG TCGGAGTATT | 16800 |
| GCATCAAGAA AATAAGGGGG GGGCAGCACA AGCTAAAAAT ATGGGGATTA GTGTAGCTAA | 16860 |
| GGGAGAGTAC ATCACGATTG TTGATTCAGA TGATATCGTA AAAGAAAATA TGATTGAAAC | 16920 |
| TCTTTATCAG CAAGTCCAAG AAAAGGATGC AGATGTTGTT ATAGGGAATT ACTATAATTA | 16980 |
| TGACGAAAGT GACGGGAATT TTTATTTTGA TGTAACAGGG CAAGATTTTT GCGTCGAAGA | 17040 |
| ATTAGCTATA CAAGAAATTA TGAACCGTCA AGCAGGAGAT TGGAAATTC AATAGCTCGGC | 17100 |
| CTTTATATTG CCGACATTTA AGTTGATTAA AAAAGAATTA TTCAATGAAG TTCACTTTTC | 17160 |
| AAATGGTCGC CGCTTTGATG ATGAAGCAAC TATGCATCGC TTTTATCTTT TAGCCTCTAA | 17220 |
| AATCGTCTTT ATAAACGATA ATCTCTATCT GTATAGAAGA CGTTCAGGAA GCATCATGAG | 17280 |
| AACGGAATTT GATCTTTCCT GGGCAAGAGA TATTGTTGAA GTGTTTCTA AGAAAATATC | 17340 |
| GGATTGTGTC TTGGCTGGTT TGGATGTCTC CGTTCTGCGT ATTCTGATTG TCAATCTTTT | 17400 |
| AAAAGATTAT AAGCAAACCT TAGAATACCA TCAATTAACA GATACTGAGG AATATAAAGA | 17460 |
| TATTTGTTTC AGATTAAAGT TGTTTTTTGA TGCAGAACA AGAAATGGTA AAAGTTGAAA | 17520 |
| TAAAGAATT GTTATTTACC ATATCACAAA CAATGAAGGT GAGGGGAGTG TTTTATGACT | 17580 |
| AAGATTTATT CGTCAATAGC AGTAAAAAAA GGACTATTTA CCTCATTTCT ACTGTTTATC | 17640 |
| TATGTATTGG GAAGTCGTAT TATCTCCCT TTTGTTGACC TAAATACTAA AGATTTTTTA | 17700 |

585

| | |
|--|-------|
| GGAGGTTCAA CAGCCTATCT AGCCTTCTCA GCCGCCCTAA CAGGTGGGAA TCTAAGAAGT | 17760 |
| TTATCAATTT TTTCTGTTGG ATTATCCCCT TGGATGTCCG CCATGATTTT ATGGCAGATG | 17820 |
| TTTTCTTTTT CTAAACGGTT GGGTTTAAAC TCTACGTCTA TAGAAATACA AGATCGCCGT | 17880 |
| AAAATGTACC TGACCTTGCT AATTGCTGTG ATTCAATCCT TGGCAGTTAG CTTGAGACTG | 17940 |
| CCAGTACAAT CCTCCTATTC TGCAATATTG GTTGTCTTAA TGAATACAAT ATTGCTGATA | 18000 |
| GCAGGAACAT TTTTCTTGT TTGGTTGTCA GATTTAAATG CGAGTATGGG GATTGGAGGT | 18060 |
| TCTATTGTAA TCCTCCTATC CAGTATGGTT TTAAATATTC CTCAGGATGT TTTGGAAACA | 18120 |
| TTTCAGACAG TACACATTCC AACAGGGATT ATTGTGTTAC TTGCTTTATT AACCCTTGTC | 18180 |
| TTTTCTTATT TACTTGCCCT TATGTATCGA GCTCGCTATT TGGTTCCTGT TAATAAAATT | 18240 |
| GGCTTACACA ATCGATTTAA ACGCTATTCT TATCTCGAAA TCATGTTGAA TCCTGCAGGT | 18300 |
| GGGATGCCTT ATATGTATGT GATGAGTTT CTTAGTGTAC CAGCTTATTT GTTCATCTTG | 18360 |
| TTGGGATTTA TTTTCCCTAA TCATTTCAGGG TTAGCGGCTT TATCAAAGGA ATTTATGGTT | 18420 |
| GGAAAGCCTT TGTGGGTCTA TGTTTATATT TCGGTCTTAT TTTTATTTAG TATCATTTTT | 18480 |
| GCTTTTGTTA CGATGAATGG AGAAGAGATT GCAGACCGTA TGAAAAATC TGGAGAATAC | 18540 |
| ATTTATGGTA TTTATCCAGG TGCGGATACT AGTCGATTTA TTAATCGATT GGTCTTCGT | 18600 |
| TTCTCAGTCA TAGGTGTCT CTTAATGTG ATTATGGCAG GTGGTCCCAT GCTTTTTGTT | 18660 |
| TTGTTTGATG AAAAGTTATT ACGATTGGCA ATGATTCCAG GCTTATTTAT GATGTTCCGG | 18720 |
| GGCATGATTT TTACGATTAG AGACGAGGTC AAGGCTTTAA GGCTAAATGA GACCTATAGA | 18780 |
| CCTTTGATTT AGGAGACTTT TATGTATTAT TTTATCCAG CTTGGTATGG GTCAGAAAGA | 18840 |
| ACATGGCATG CAGATATCAC TCCATGGTAT TTTCTCTATT TTCGTCTAGA GTTTGATGAT | 18900 |
| ACCTTTCACC AGATTCGGCT CTTTCAAGAG CAAGATATAG ATTCTCGTCT ATTAGTATTA | 18960 |
| GCTTACCAGC CTCATCTACG TTATTTTTTA TATAGACATG GTGTGTTAGA AATGGATACT | 19020 |
| TATTCCGTTT TTGATGTAT GCAAGATTT CATAATCTCC ATACCCAAGT TTTAAGCATT | 19080 |
| AGAGATATTG AGTGGGATGA TGA CTGTGAA TTTATTTATA GTCCCTTTAC GATTATCGTT | 19140 |
| CAAAAAATG GGAAGAAATT TGCTAAGGTT GAACATGGAG TTGAAGGCTT CATCAGTGAT | 19200 |
| ATACAGTATT TTGAACCAAA TGGTCAAATA CATATGCACC ATATCGTGGA TGATCGTGGG | 19260 |
| TTTGTATCGA GCATTATCTT TTTTGAAGAT GGGCAAGCAG CCTATCAAGA ATATCTGAAC | 19320 |
| CTCAAGGGAG AGTGGCAATT CAGAGAGCGT TTAAAAGAAG GAGGACAGGT AGAAGTCAAT | 19380 |
| CCAATTTTGG GTTATCGCTT TAAATGCTT ACCTATCAAA ATATGGGAGA TCTGGTGGCA | 19440 |

586

| | |
|--|-------|
| GAATTTTTTG AGAATTATCT GCAAACGTAT GTGAAGGATC AGGATATTTT TATGCTTCCT | 19500 |
| TCTCATTCTC ATCATGACCA GTTGGTACTA GATCGTTTAC CTAGTACTAA TCCTAAACTG | 19560 |
| TTGAGTCTGT TCATTGGACG TAATCCTCAA GATACCTTTA GGGATTTAGA TGTAACTTTT | 19620 |
| GAAAAATCGG ATTTGATTTT GGTGGATAGA GAGGATAGTT TACGATTGTT GCAGGAGTTG | 19680 |
| TATCCTGAAC GAATGCATCA ATGTTATCAT TTATCATCTT TTGACACCCG ATTACGATTG | 19740 |
| GGACGAAGCC AAAC TAAGAA AGAATCCATC ATTTATTTTC AACTGGATTT TGAGCAGGGG | 19800 |
| ATTGATAATC AAGCTCTGCT TCAAGTCTTG TCCTTTGTCG CTGAAAATAA GGATACTGAG | 19860 |
| GTGATTTTTG GAGCCTTTGC TGCTAGTCAG GAGCAAATGA ATGAGGTTGA AGGGATTGTT | 19920 |
| GAGTCTTTCA TCCAAGAAAA CATTC AATCC GAAATCTGG GAAAGGCGAT TGATTATGGT | 19980 |
| GATGCAGAAA ATCCTCTGGA AGAAAAATA CACCAGGACT TACGCTTACA GTTTGTTAAC | 20040 |
| TTGAATGATG AGTTAGATTT GATAAAAACA CTAGAATTTG TCCGTTTGAT TGTGGATTTA | 20100 |
| AATAGACATC CTCATCTCTA CACACAGATT GCTGGGATTA GTGCAGGAAT TCCTCAAATC | 20160 |
| AACCTAGTTG AAACCGTCTA TGTGTAACAT TTAAAAATG GTTATTTGTT AGCAGATGTT | 20220 |
| ACAGAAATTT CTAAGGCTGC ACATTATTAC ACAGATAGGT TGAAGGAGTG GAATGAGTCC | 20280 |
| TTGATATATT CAATTGATAA GATTAAGGAG CACACAGGAC AACAATTTCT TGGAAAATTA | 20340 |
| GAGAAATGGA TAGAGGAGGT TAAAAATGTC AAAGGAACTT AATATTTTAC AGATAGGACT | 20400 |
| TGCCAATTGG GAAAACTACT ATGACATACC TGAAAAATG AGTTGGTATT ATTTTACCC | 20460 |
| AAACTCATCA AAAGCCCTTC GTGAAATAAT TGAAAAAGAG GATATTAACC GTTTTCATGC | 20520 |
| AGTTTTTAATA GAAGATGGTC AGTATCCAG AGACTTATTT TCCTATGTAA AATATTTTGA | 20580 |
| ACCTTATACT TTATTTTATA ACCAGAATCT ACAAATAAAT GATAGAGAGG TTGTGGATTT | 20640 |
| TCTAAAAAAA CGATGTGCAC AAGCAATTGA TTTTTTAAGT CCCCACAAC TAATCAATGA | 20700 |
| TTTAAGTAAA TCTCTTTTGG GCGGTGGGTA TGGTGATAAA CTCTTTCCTC CGACGATACA | 20760 |
| AGTCAATCCA AATTTTACAG GAGCTATTTT TTATCAAGGA TTGGATTATG TCAGTTTGGA | 20820 |
| AGGTGAGTTT GGGCAAGATT TTGCCCAGCT TGCCTATTGG GCTTATAATA TTATGGTGCA | 20880 |
| AAAAACACTC CCTATTGAGT TGTGGCTTGA ATATGAGAAG GAAGGCAATT GTGACTTTCG | 20940 |
| TTTAGTAATC CGTAAAAATG GGAGTGGGTC TGTTGATGAT TTCTTTGAAG AAGTAATAGT | 21000 |
| ATCTGAAAAA GACTTGGAGC AAGCACTTTT TATGGATAGT CGAGACGGAG ACTACTTTCT | 21060 |
| CTCGATATCT GTTGAAGCAA GAGGTCGTGG AACTATCAAA CTAGGTAATC TTCACCAACG | 21120 |
| ATGGAGTCGA AAACAATTTG GTAAGTTTGT ACTTGGTGGA AATATCCTAC ATGATTCCAA | 21180 |
| GCGTGATGAA ATAAACTATT TCTTCCATCC AGGTGATTTT AAACCGCCTT TGACTGTCTA | 21240 |

587

TTTTCGAGGT TATCGACCTG CAGAAGGATT CGAGGGTTAC TTTATGATGA AAACCTCTGG 21300
ATGTCCCTTC ATTTTATTTT CTGATCCACG TTTAGAGGGG GGAGCTTTTT ATCTCGGAAC 21360
GGATGAGCTA GAGGGAAAAG TAAAGGATAC GATCACTCAC TATCTTGATT ATTTAGGCTT 21420
TGATCATAAG GATTTGATTT TATCAGGTCT TTCTATGGGA ACGTTTCCGG CTCTCTATTA 21480
TGGTGCTTCT TTTGAACCCC ATGCCATCAT AGTTGGTAAG CCCTTGGCTA ATTTAGGAAC 21540
TATAGCTAGT CGTGGACGTT TGGACGCACC GGGTGTCTCT AACTTAGCTT TTGATTGTTT 21600
AATTCATCAT ACAGGTGGGA CAAGTTCTCA AGATATGACG GAGTTGGATC AGCGTTTTTG 21660
GAAAATTTT AAACAAGCAA ATTTTCAAA GACAACCTT GGTATATCCT ATATGAAAGA 21720
TGAAGAAATG GATCCACAAG CCTATGAACA ATTAGTGTCT TATCTGTGTA ATACAGGTGC 21780
GAAGATTTTA TCTAAAGGAA CTGCTGGACG ACACAATGAT GATACAGATA CCAATATTTT 21840
TTGGTTTTTG CACTTTTATA GAATGGTCTT AGAGACTGGT TTTGGAAGGG AGAAAAGATG 21900
ATTATTACAC AGAGACAGTC TATTCATTGG GGAGAAGTTG GTGGGACTTA TATGTATGGA 21960
ACAACTGTAT CTTATTACCC TGACAAAAGT GTTCGTCTGT ATAATCCTCT ATTGCCATCT 22020
GGTGAGATTC TAAAGACTTG GTTTTCTAGT GTCAATTACC AGGCTGCACG AACCCAACCT 22080
CAGCTTCCCT TATTAAAAAG AAAGCAGGAG TATCAACTAT CACTGGTTTT TGACTGTCAG 22140
CCTGAAAATG GAGTTTATAC CAAGATAACT TTTTTTGACC GCTATGGTGA TATTTTAGAA 22200
AAAAAGGTAG AAAAAGTGAA AGATTTCATA TTTACTTATC CAGAAGATAG TTATACTTAT 22260
CGAGTTTCTC TTTTAAGTGC TGGATTGAG TCCTTAACTT TTTATCATTT TTCTATCAAG 22320
GAGATCAGAA GTGTTTAGAC GTTTAGGTCA AGATTTCAG CTTAGGAAAG TGAAAAGAT 22380
TTTAAAGCAG ATTAATGCCC TGAAAGGCAA GATGTCCTCT CTTTCGGATC AAGAATTAGT 22440
AGCTAAAACA GTAGAGTTTC GTCAGCGTCT TTCCGAGGGA GAAAGTCTAG ACGATATTTT 22500
GTTTGAAGCT TTTGCTGTGG TGCGTGAAGC AGATAAGCGG ATTTTAGGGA TGTTTCCTTA 22560
TGATGTTCAA GTCATGGGAG CTATTGTCAT GCACTATGGA AATGTGCTG AGATGAATAC 22620
GGGGGAAGGT AAGACCTTGA CAGCTACCAT GCCTGTCTAT TTGAACGCTT TTTCAGGAGA 22680
AGGAGTGATG GTTGTGACTC CTAATGAGTA TTTATCAAAG CGTGATGCCG AGGAAATGGG 22740
TCAAGTTTAT CGTTTCTAG GATTGACCAT TGGTGTACCA TTTACGGAAG ATCCAAAGAA 22800
GGAGATGAAA GCTGAAGAAA AGAAGCTTAT CTATGCTTCG GATATCATCT ACACAACCAA 22860
TAGTAATTTA GGTTTTGATT ATCTAAATGA TAACCTAGCC TCGAATGAAG AAGGTAAGTT 22920
TTTACGACCG TTAACTATG TGATTATTGA TGAAATTGAT GATATCTTGC TTGATAGTGC 22980

| | | | | | | |
|------------|------------|------------|--------------|-------------|------------|-------|
| 588 | | | | | | |
| ACAAACTCCT | CTGATTATTG | CGGGTTCTCC | TCGTGTTTCCAG | TCTAATTACT | ATGCGATCAT | 23040 |
| TGATACACTT | GTAACAACCT | TGGTCGAAGG | AGAGGATTAT | ATCTTTTAAAG | AGGAGAAAGA | 23100 |
| GGAGGTTTGG | CTCACTACTA | AGGGGGCCAA | GTCTGCTGAG | AATTTCTCTAG | GGATTGATAA | 23160 |
| TTTATACAAG | GAAGAGCATG | CGTCTTTTGC | TCGTCATTTG | GTTTATGCGA | TTCGAGCTCA | 23220 |
| TAAGCTCTTT | ACTAAAGATA | AGGACTATAT | CATTTCGTGGA | AATGAGATGG | TACTGGTTGA | 23280 |
| TAAGGGAACA | GGGCGTCTAA | TGGAAATGAC | TAAACTTCAA | GGAGGTCTCC | ATCAGGCTAT | 23340 |
| TGAAGCCAAG | GAACATGTCA | AATTATCTCC | TGAGACGCGG | GCTATGGCCT | CGATCACCTA | 23400 |
| TCAGAGTCTT | TTTAAGATGT | TTAATAAGAT | ATCTGGTATG | ACAGGGACAG | GTAAGGTCGC | 23460 |
| GGAAAAAGAG | TTTATTGAAA | CTTACAATAT | GTCTGTAGTA | CGCATTCCAA | CCAATCGTCC | 23520 |
| GAGACAACGG | ATTGACTATC | CAGATAATCT | ATATATCACT | TTACCTGAAA | AAGTGTATGC | 23580 |
| ATCCTTGGAG | TACATCAAGC | AATACCATGC | TAAGGGAAAT | CCTTTACTCG | TTTTTGTAGG | 23640 |
| CTCAGTTGAA | ATGTCTCAAC | TCTATTCGTC | TCTCTTGTTT | CGTGAAGGGA | TTGCCATAA | 23700 |
| TGTCCTAAAT | GCTAATAATG | CGGCGCGTGA | GGCTCAGATT | ATCTCCGAGT | CAGGTCAGAT | 23760 |
| GGGGGCTGTG | ACAGTGGCTA | CCTCTATGGC | AGGACGTGGT | ACGGATATCA | AGCTTGGTAA | 23820 |
| AGGAGTCGCA | GAGCTTGGGG | GCTTGATTGT | TATTTGGACT | GAGCGGATGG | AAAGTCAGCG | 23880 |
| GATCGACCTA | CAAATTCGTG | GCCGTTCTGG | TCGTCAGGGA | GATCCTGGTA | TGAGTAAATT | 23940 |
| TTTTGTATCC | TTAGAGGATG | ATGTTATCAA | GAAATTTGGT | CCATCTTGGG | TGCATAAAAA | 24000 |
| GTACAAAGAC | TATCAGGTTT | AAGATATGAC | TCAACCGGAA | GTATTGAAAG | GTCGTAAATA | 24060 |
| CCGGAAACTA | GTCGAAAAGG | CTCAGCATGC | CAGTGATAGT | GCTGGACGTT | CAGCACGTCG | 24120 |
| TCAGACTCTG | GAGTATGCTG | AAAGTATGAA | TATACAACGG | GATATAGTCT | ATAAAGAGAG | 24180 |
| AAATCGTCTA | ATAGATGGTT | CTCGTGACTT | AGAGGATGTT | GTTGTGGATA | TCATTGAGAG | 24240 |
| ATATACAGAA | GAGGTAGCGG | CTGATCACTA | TGCTAGTCGT | GAATTATTGT | TTCACTTTAT | 24300 |
| TGTGACCAAT | ATTAGTTTTC | ATGTTAAAGA | GGTTCAGAT | TATATAGATG | TAACTGACAA | 24360 |
| AACTGCAGTT | CGTAGCTTTA | TGAAGCAGGT | GATTGATAAA | GAACTTTCTG | AAAAGAAAGA | 24420 |
| ATTACTTAAT | CAACATGACT | TATATGAACA | GTTTTTACGA | CTTTCAGTGC | TTAAAGCCAT | 24480 |
| TGATGACAAC | TGGGTAGAGC | AGGTAGACTA | TCTACAACAG | CTATCCATGG | CTATCGGTGG | 24540 |
| TCAATCTGCT | AGTCAGAAAA | ATCCAATCGT | AGAGTACTAT | CAAGAAGCCT | ACGCGGGCTT | 24600 |
| TGAAGCTATG | AAAGAACAGA | TTCATGCGGA | TATGGTTCGT | AATCTCCTGA | TGGGGCTGGT | 24660 |
| TGAGGTCACT | CCAAAAGGTG | AAATCGTGAC | TCATTTTCCA | TAAAAGGAGA | AAATATGACA | 24720 |
| ATTTACAATA | TAAATTTAGG | AATTGGTTGG | GCTAGTAGCG | GTGTTGAATA | CGCTCAAGCC | 24780 |

589

| | |
|---|-------|
| TATCGTGCTG GTGTTTTTCG GAAATTAAAT CTGTCCTCTA AGTTTATCTT TACAGATATG | 24840 |
| ATTTTAGCCG ATAATATTCA GCACTTAACA GCCAATATTG GTTTTGATGA TAATCAGGTT | 24900 |
| ATCTGGCTTT ATAATCATTT CACAGATATC AAAATTGCAC CTA TAGCGT GACAGTGGAT | 24960 |
| GATGTCTTGG CTTACTTTGG TGGTGAAGAA AGTCACAGAG AAAAAATGG CAAGGTTTTA | 25020 |
| CGTGTATTCT TTTTGACCA AGATAAGTTT GTAACCTGTT ATTTGGTTGA TGAGAACAAG | 25080 |
| GACTTGGTTC AACATGCCGA GTATGTTTTT AAGGGAAACC TGATTCGGAA GGATTACTTT | 25140 |
| TCTTATACGC GTTATTGTAG CGAGTATTTT GCTCCCAAGG ACAATGTTGC AGTCTTATAC | 25200 |
| CAACGAACCT TTTATAATGA AGACGGGACT CCAGTCTATG ATATCTTGAT GAATCAAGGG | 25260 |
| AAGGAAGAAG TTTATCATTT CAAGGATAAG ATTTTCTATG GAAAGCAAGC TTTTGTGCGT | 25320 |
| GCCTTTATGA AATCTTTGAA TTTGAATAAG TCTGATTTGG TCATTCTCGA TAGGGAGACA | 25380 |
| GGTATGGAC AGGTGTGTT TGAGGAAGCA CAGACAGCAC ATCTAGCGGT AGTTGTTTAT | 25440 |
| GCGGAGCATT ATAGTGAAAA TGCTACAAAT GAGGACTATA TCCTTTGGAA TAACTATTAT | 25500 |
| GACTATCAGT TTACCAATGC AGATAAGGTT GACTTCTTTA TCGTGTCTAC TGATAGACAA | 25560 |
| AATGAAGTTC TACAAGAGCA ATTTGCCAAA TATACTCAGC ATCAGCCAAA GATTGTTACC | 25620 |
| ATTCCTGTAG GCAGTATTGA TTCCTTGACA GATTCAAGTC AAGGGCGCAA ACCATTTTCA | 25680 |
| TTGATTACGG CTTACGCTCT TGCCAAAGAA AAGCACATTG ATTGGCTTGT GAAAGCTGTG | 25740 |
| ATTGAAGCTC ATAAGGAGTT ACCGGAACCT ACCTTTGATA TCTATGGTAG TGGTGAGAA | 25800 |
| GATTCTCTGC TTAGAGAAAT TATTGCAAAT CATCAGGCAG AGGACTATAT CCAACTCAAG | 25860 |
| GGGCATGCGG AACTTTCGCA GATTTATAGC CAGTATGAGG TCTACTTAAC GGCTTCTACC | 25920 |
| AGCGAAGGAT TTGGTCTGAC CTTGATGGAA GCTATTGGTT CAGGTCTACC TCTAATTGGT | 25980 |
| TTTGATGTGC CTTATGGTAA TCAGACCTTT ATAGAGGATG GGCAAAATGG TTATTTGATT | 26040 |
| CCAAGTTCAT CTGACCATGT AGAAGACCAA ATCAAGCAAG CTTATGCCGC TAAGATTTGT | 26100 |
| CAATTGTATC AAGAAAATCG TTTGGAAGCT ATGCGTGCCT ATTCTTACCA AATTGCAGAA | 26160 |
| GGCTTCTTGA CCAAAGAAAT TTTAGAAAAG TGGAAGAAAA CAGTAGAGGA GGTGCTCCAT | 26220 |
| GATTGAACTT TATGATAGTT ACAGTCAAGA AAGTCGAGAT TTACATGAAA GTCTAGGCGC | 26280 |
| TACTGGTCTT TCTCAACTTG GAGTGGTCAT CGATGCAGAT GGTTTTCTGC CTGATGGTCT | 26340 |
| GCTTTCTCCT TTTACCTATT ATCTAGGTTA CGAGGATGGA AAACCTCTCT ATTTTAATCA | 26400 |
| AGTTCCCGTT TCAGATTTTT GGGAAATTTT AGGAGATAAT CAGTCTGCTT GTATTGAAGA | 26460 |
| TGTGACGCAG GAGAGGGCTG TCATTCATTA TGCTGATGGA ATGCAGGCTC GCTTGGTTAA | 26520 |

590

| | |
|---|-------|
| ACAGGTAGAC TGGAAAGACC TAGAAGGTCG AGTACGTCAG GTTGACCACT ACAATCGCTT | 26580 |
| CGGAGCTTGT TTTGCTACAA CGACTTATAG CGCAGATAGC GAGCCGATTA TGACAGTTTA | 26640 |
| CCAAGATGTC AATGGTCAAC AAGTTTACT GGAAAACCAT GTGACGGGTG ATATCTTATT | 26700 |
| GACTTTGCCA GGTCAGTCCA TGCCTTACTT TGCAAATAAA GTTGAATTTA TCACCTTCTT | 26760 |
| TTTGCAAGAT TTGAAATAG ATACCAGTCA GCTTATCTTT AATACTCTAG CGACTCCTTT | 26820 |
| CTTGCTTTCC TTCCATCATC CAGATAAATC TGGCTCGGAT GTCTTGGTAT GGCAGGAACC | 26880 |
| TCTCTATGAT GCCATTCCAG GTAATATGCA GTTGATTTTG GAAAGTGATA ATGTGCGTAC | 26940 |
| TAAGAAGATC ATCATTCCAA ATAAGGCGAC TTATGAGCGC GCTTTAGAGT TAACTGACGA | 27000 |
| GAAATACCAT GATCAGTTTG TGCACTTGGG TTATCATTAC CAGTTCAAAC GTGATAATTT | 27060 |
| CCTAAGACGA GATGCCTTAA TCTTGACCAA TTCAGATCAG ATTGAGCAAG TAGAAGCAAT | 27120 |
| CGCAGGAGCC TTGCCTGATG TCACTTTCCG TATTGCAGCG GTGACAGAGA TGTCTTCTAA | 27180 |
| GCTCTTAGAC ATGCTTTGCT ATCCTAATGT GGCCCTTTAC CAGAACGCTA GTCCACAGAA | 27240 |
| GATTCAGGAG CTGTATCAAC TGTCGGATAT TTAATTGGAT ATAAACCACA GTAATGAGTT | 27300 |
| GCTACAGGCA GTGCGTCAGG CCTTTGAGCA CAATCTCTTG ATTCTTGGCT TTAATCAGAC | 27360 |
| GGTGCACAAT AGACTTTATA TCGCTCCAGA CCATCTATTT GAAAGTAGTG AAGTTGCTGC | 27420 |
| TTTGGTTGAG ACCATTAAAT TGGCCCTTTC AGATGTTGAT CAAATGCGTC AGGCACTTGG | 27480 |
| CAAACAAGGC CAACATGCAA ATTATGTTGA CTTGGTGAGA TATCAGGAAA CCATGCAAAC | 27540 |
| TGTTTTAGGA GGCTAACATG TCAGAGGAAG ATTTATTTTA CAAAGACGTT GAAGGCCGCA | 27600 |
| TGGAAGAGTT GAAACAAAAA CCCATCAAGA AGGAAAAAGA AAGCCGAGGG GAAAGATTA | 27660 |
| GTAAGACTTT TTCACTTTTA CTGGGTTTGA TGATTCTGAT TGTTTGGCTC TTTACTTTGC | 27720 |
| TGGAATTTT GAGGTAGATC TATGATTGAA ATACTAATTG TTTTAGCTAT TATCCTATCT | 27780 |
| CTTGCTTTGA TTGTATTGGT AACTATACAA CCCCCTCAA ATCAACTATT TTCCATGGAT | 27840 |
| GCCACTAGTA ATATTGGTAA ACCAAGCTAC TGGCAGAGCA ACACCTTGGT CAAGGTGCTC | 27900 |
| ACTTTATTGG TGAGTTTGGC TTTATTTATT CTACTATTAA CCTTTATGGT GATTACTTAT | 27960 |
| AAATAAAAGA AAATTCAGA TATTCACCTT TTGTGGATTG GTCTGAAGTT TTCTTTTTTA | 28020 |
| TACTCAATGA AAATCAAAGA GCAAACTAGG AAGCTAGCCG CAGGCTGCTC AAAACACCGT | 28080 |
| TTTGAGGTTG TAGATATAAC TGACGAAGTC AGCTCAAAAC ACCGTTTGA GGTGTAGAT | 28140 |
| ATAACTGACG AAGTCAGCTC AAAACACCGT TTTGAGGTTG TGGATAGAAC TGACGAAGTC | 28200 |
| AGCTCAAAAC ACCGTTTGA GGTGTGGAT AGAACTGACG AAGTCAGCTC AAAACACCGT | 28260 |
| TTTGAGGTTG TGGATAGAAC TGACGAAGTC AGCTCAAAAC ACCGTTTGA GGTGTGGAT | 28320 |

591

| | |
|---|-------|
| AGAACTGACG AAGctCAGTA ACATATATAC AGCAAGGCGA CGCTGACGTG GTTTGAAGAG | 28380 |
| TATTACTGTC TATATTTTGT GTAAAAATCA ACTTTTACTT GGATGAAGGT TTTGGCTTCA | 28440 |
| CGTAGGAGTT GAAGAAGGGT GGCGCGGGTT TCAAATTCCTT CTCTTGCTCTT GGGCAGACTG | 28500 |
| CGGTTCGGGA AGACTTCCAG ATAACGTTCA ATTTTCATCTA GCAAATCAGA AGCAGGATTG | 28560 |
| GTCTGGCTCA GTTGACCTGC AATTTTGTGA AAGAGTTGCG CTAAGATCAG GCTTTCAGTG | 28620 |
| GCGGCAAGGT GACAAGTGT AATCTGTTGG GCCATGTTTC TCAGGATACG ACTTTGTCGC | 28680 |
| TGTCTCATCT CAAAGTAGTG GATATGGTAG TCTGTCTGGT GAAAGAGGTG GTCAGAGTGA | 28740 |
| TCCAAATAGA CCAGTCTGAG GGCTTCTTTC AAAAGCGTGT CTAATTCTGC TACCAGCTGT | 28800 |
| GCTCGGTTGC GTCCGTCTCC TCTGGATAAA TAGTATTTGA AGCGCTGGAG GATATCTTTT | 28860 |
| AACTTTTCTT CCACCAGCGT GTGGTAGTGC TGGATTTTCTT CTTCTCGTGA AGGCATATAG | 28920 |
| AGATTAACAA GCAAGGCAAA TCCTGTACCA ATAGCAAAGA GAAGGAATTC ATTGACTAGA | 28980 |
| AGGTCTGGAG AGGTTGACTC TTGAACCAAG AGATGGCTAA CCAAAACAGT GCTTGGTGTG | 29040 |
| ATGCCAATTT CCCAGCCCAT CTTGTAGGCT AAAGGAACGT AGAAGGCCAG ATAGAGGCCG | 29100 |
| AGACTCCAGA TATGAAATCC GCTCAAGTGA AAAGCTAGAA CACCGATAGC CAGAGCTAGA | 29160 |
| AGCATAGAAA AAAGACGATT GCGAGCCAGT TTTAAAGTAC TTCTACGCGT ATCAGATAGG | 29220 |
| CTCAAGAGAG CGATAATTCC AGCCGAAACT GCTGACGAAA GATTGAGAAA ATAAGCAAGC | 29280 |
| AGGCAGGCAA GACAGGTAGC TAAGATGAGC TTGGTCGTAC GTTGGCTAAT AGACATAAGA | 29340 |
| ATTTCTTAAT AAGTTAGAAT AAAAGCGTAA AAGACAAGAC ATGAGCAGGC TTGCCTTGAT | 29400 |
| GAGTTATTTT TTACGGGTTG CTGCGTATTC GGCAACGGCG GTAAAGAGGA CATCTGTAGA | 29460 |
| AGAGTTAAGG GCTGTTTCAC ATGAGTCTTG GATGACACCA ATCACAACAA CAACCCCAAC | 29520 |
| AATTTGTATG GCAATATCGT TAGAAATACC GAAAAGGCTA CAAGCAACTG GGATAAGAAG | 29580 |
| GAGGGAACCT CCGGCAATAC CTGAAGCATC ACAGGATGAG ATAGCTGCTA CCACACTGAG | 29640 |
| GACAAAGGCT GTGGCAAAGT CAACAGGAAT TCCAAGAGTG TTAAGTGCAG CAAGGGTCAA | 29700 |
| AAGGTTAATG GTAATCGCTA CTCCAGCCAT ATTGATAGTA GAACCGAGTG GGATAGAAAC | 29760 |
| AGAATAGGTA TCTGGGTGTA GTCCAAGGTC ATGGCAGAGT TTCATGTTGA CAGGAATGTT | 29820 |
| AGTCGCAGAA CTACGAGTGA AAAAGGCTGT CACACCGCTG ACACGGAGGC AGTTCCAAAC | 29880 |
| TAGAGGGTAA GGATTGCGTC TCATAAAGAA GAAGGCAATC AAAGGGTTGA CCACAGGGGC | 29940 |
| AACAAAAAGC ATAGTCGTTA CTAATAGAAC CAATAAAATA CCGTAGTTGG CAAGGCTTCC | 30000 |
| GACTCCCTTG TCAGAAATGG TTTTAAAAAC AAGACCAAGG ATTCCAAATG GAGCCAGATT | 30060 |

592

| | |
|--|-------|
| GATGATCCAT TCGACAATTT TAGAAGTCAC GTCAGCGATA GTTTTTAGCA ATTCTTGACT | 30120 |
| ATTTTTACTG GCTTCTCTCA TAGCGATTCC AAAAATGACT GCCCAAGATA AGATTCTAAT | 30180 |
| ATAGTTAGCA GTAAGCAGGG CGTTGACTGG GTTGTCACC AGTTTGAGCA AGAGGTGCT | 30240 |
| GAGAACCTGC CCAATCCCAT CTGGTGGTGC AATTTCAGTA TTGGCACTAT TTGGGGTAAT | 30300 |
| TTCAATAGGG ACGATGAAAT TTGCTAGTAC AGCTACAAGA GCAGCGGCGA AAGTCCCTAT | 30360 |
| CATAGGATAT ACAAGAAAAC AACAGTTTTT ATATTGCTAT CTTGTCCCTT TTGATGTTGG | 30420 |
| GAAAGGGCAT TGGCAACGAG AGCAAAGACT AGGATAGGAG CAACAGCTTT TAGACCTCCA | 30480 |
| ACGAATAAAT CCTCGAGTAG CCCAATCCCT GAGAGATTAG GAAGGGTCAG TCCTAGGATT | 30540 |
| CCCCACAAAG CATACCAATC AAGATACGCT TGACAAGGCT TGCCCTATTTC CAAGCATGAA | 30600 |
| TGATTCTTTT CATAATAATC TCCTTTTGT GTAGTGATTA TGATTATAGT ATAAATGATA | 30660 |
| GACAAAATCA AGAATTTTCT GTCTATTTTT TGAATATTTA TGGAGAATGA GACTGATGAA | 30720 |
| AATATGGTAT AATGAAATAA AGGAGTTTTA TATGCAAAA TTTATTCAGG CTTATATTGA | 30780 |
| AAAGCTAGAT GTGACAACCA TTATCGAGAA TATTCTAACC AAGGTCATTT CTCTTTTACT | 30840 |
| GCTTTTAATT GTATTTTATA TTGCTAAAAA AATGCTTCAT ACCATGGTGC AGAGAATGT | 30900 |
| CAAACCTTCT CTAATAATGT CTCGTCATGA TGTGGACGC CAAAAAACCA TCTCACGTTT | 30960 |
| ACTAGAAAAT GTGTTTAATT ATACGCTATA TTTCTTTTTA CTCTACTGCA TTTTGTGAT | 31020 |
| TTTAGGTTTG CCAGTTTCTA GTTTGCTGGC TGGAGCTGGT ATTGCTGGGG TAGCGATTGG | 31080 |
| TATGGGAGCC CAAGGCTTTC TGTCTGATGT CATCAATGGC TTTTTCATCC TCTTTGAACG | 31140 |
| TCAACTGGAT GTGGGAGATG AGGTCGTTCT GACAAATGGA CCGATTACTG TATCGGGTAA | 31200 |
| GGTTGTCAGT GTGGGAATTC GTACGACACA GCTTCGTAGC GAGGAGCAAG CCCTTCACTT | 31260 |
| TGTCCCTAAC CGAAATATCA CAGTTGTTAG CAATTTCTCA CGCACAGACT AGACCTGTTA | 31320 |
| TTTTAAGTAA TTTGTGGTAC AATAGAGGGA GTTTAATAAG GAGAAAAGAT GGTTTTAGAA | 31380 |
| AAGCAGTTGG GCAATGGTTG TACCTGGATA GACCTAGACC TAGGAAAGTT GAATAAATA | 31440 |
| GAAGACCTTT CTGAAATTTA CGGTTTGGAC AAGGAAACCA TTGAATACGC ACTGGATAGA | 31500 |
| AACGAGCGCG CCCACATGGA CTACCACCGT GAAAGTGAGA CGGTTACCTT TATCTATAAT | 31560 |
| GTCTTAGACG TAAAAAGGA CAAGGCCTAC TATGAGACTT TTCCCATGAC CTTTATTGTC | 31620 |
| GAGCATCGTC GCCTGATTAC CATTAGTAAT ACCAAGAACG CCTATGTCAT TGAACAGATG | 31680 |
| ACTCGTTATC TGGAGAACCA TGACACGCTT TCGATTTATA AGTTTCTCTT TGCCAGTCTG | 31740 |
| GAAATCATCA GCAATGCCTA CTATCCTGTC ATTGAGCAGA TGGACAAGAG TAGGGATGAG | 31800 |
| GTCAATGACC TCTTGCGCCA GCGAACTACC AAGAAAAACC TCTTTGTCCT GTCTGATTGG | 31860 |

593

| | |
|---|-------|
| GAGACTGGTA TGGTTTATCT GACGGCAGCT GCCAAACAAA ATCGGATTTT GTTAGAGCAT | 31920 |
| ATCAAGGTC ATGCCTTGTA TCGTAGTTTT GATGAGATTG AGAGAGAACA GTTTGATGAT | 31980 |
| GCCATGATTG AGGCTCATCA GCTGGTATCC ATGACAGACC TAATCTCTCA GATTTTACAG | 32040 |
| CAGCTTTTCTAG CCTCTTACAA CAATATTCTA AACAATAATC TGAATGACAA TTTGACAACC | 32100 |
| TTGACTATCA TTTCAGTCTT GCTAGCTGTT TTGGCAGTCG TGACAGGCTT TTTCGGAATG | 32160 |
| AATGTTCCCT TACCTTTAAC AGATGAGCCC CATGCTTGGC TCTATATCAG TTTGGCTAGT | 32220 |
| GCAGGTTTGT GGATTGTTTT ATCCTTGTTA CTAAGGAAAA TTGCGAAAAA AAGTTAAGAA | 32280 |
| AAGGAGCCAG AATGGCGATT GAAAATTATA TACCAGATTT TGCTGTGGAA GCAGTCTATG | 32340 |
| ATCTGACAGT CCCAAGCCTG CAGGCGCAGG GAATAAAGGC TGTTTTGGTC GATTGAGATA | 32400 |
| ATACCCCTCAT TGCTTGGAAC AACCTTGATG GAACGCCAGA GATGAAGCAA TGGCTACATG | 32460 |
| ACCTTCGGGA CGCGGGTATT GGCATTATCG TAGTGTCAAA TAACACCAAA AAACGCGTTC | 32520 |
| AACGAGCAGT TGAGAAATTT GGGATTGATT ACGTTTACTG GGCCTTGAAG CCCTTCACAT | 32580 |
| TTGGTATTGA CCGTGCTATG AAGGAATTCC ACTATGACAA AAAGGAAGTG GTCATGGTTG | 32640 |
| GTGACCAACT CATGACAGAT ATACGAGCAG CCCACCGTGC AGGGATTTCGG TCAATTTTAG | 32700 |
| TCAAACCCCTT GGTCCAACAT GACTCAATCA AAACGCAGAT TAACCGAACT CGTGAGCGTC | 32760 |
| GTGTTATG | 32768 |

(2) INFORMATION FOR SEQ ID NO: 72:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 14872 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

| | |
|---|-----|
| CCAGTCACAA AGAAATTGAG CGCGTTCAGC TGAGGATGCA CTATGATGCA AGCTACATTT | 60 |
| CATTTGATGG GATATTAAGA AAGGAGATTT TCATGACACT TTTAGATGTA AAACACGTTT | 120 |
| AAAAAATTTA TAAACACGTTTTCAGGGCA ACCAAGTAGA AGCCCTCAAG GATATTCACT | 180 |
| TTACCGTAGA AAAGGGTGAC TACGTTGCCA TCATGGGTGA GTCTGGTTCT GGTAAATCAA | 240 |
| CTCTTCTCAA TATTCTAGCT ATGTTGGATA AACCAAGTCG TGGTCAGGTT TACTTGAATG | 300 |
| GAACAGACAC CGCAACTATT AAAAATTCAC AGGCTTCTAG TTTCCGGCGT GAAAAGCTAG | 360 |
| GATTTGTCTT CCAAGACTTT AACTTGCTAG ATACTCTGTC TGTTAAGGAC AATATCTTGC | 420 |

594

| | |
|---|------|
| TTCCGCTTGT CTTGTCAAGA AGACCTATAA CGGAGATGAT GAAGAAATTG GTGGTGACAG | 480 |
| CTGAGAATCT GGGTATTAAC CAATTGCAAG AGAAGTACCC TTACGAGATT TCTGGTGGTC | 540 |
| AGAAACAGCG TG TAGCAGTA GCCCGCGCCA TCATCACAGA ACCTGAAATT CTCCTTGCGG | 600 |
| ACGAGCCAAC AGGAGCCCTT GATTCCAAGT CATCTGCAGC CTTACTTGAT GTCTTTAATG | 660 |
| AAATCAATGA GCGTGGGCAA ACCATCCTCA TGGTAACCCA CTCAACAGCA GCTGCTAGCA | 720 |
| GGGCCAAGCG TGTCTCTTT ATCAAAGACG GCATTCTTTA CAACCAAATC TACCGTGGAG | 780 |
| AGAAGACAGA GCGTCAGATG TTCCAAGAAA TCTCTGATAC CTTGACTGTC ATGGCAAGCG | 840 |
| AGGTGAATTA GTATGTTTCG ATTAACCAAT AAGTTAGCGG TATCGAACTT GATTA AAAAC | 900 |
| CGCAAACCTCT ACTATCCCTT TGCCTGGCT GTTCTCTTGG CAGTCACCAT CACCTATCTC | 960 |
| TTTTACTCCC TAACCTTCAA TCCAAAGATT GCGGAAATCC GTGGAGGAAC CACCATTCAA | 1020 |
| GCAACACTTG GATTTGGTAT GTTGTGCGT ACCCTGCGT CACCATTTATC GTCCTCTATG | 1080 |
| CCAATAGTTT TGT CATGAAA AACCGTTCCA AGGAACTGGG TATATATGGC ATGTTAGGCT | 1140 |
| TGGAGAAGCG CCATCTAATC AGTATGACCT TTAAGGAGTT AGTGGTATTT GGGATTCTAA | 1200 |
| CTGTTGGAGC GGGTATCGGT ATTGGAGCCT TGTTTGACAA GTTAATTTTC GCTTTCCTGC | 1260 |
| TCAAATAAT GAAACTGAAG GTTGAGCTGG TTGCTACCTT CCAAATGAAT GTTGTCAATG | 1320 |
| CAGTACTTGT TGTCTTTGGA TTGATTTTCC TAGGCCTCAT GTTCCTGAAT GCTCTTCGAA | 1380 |
| TCGCCCCGTAT GAATGCCCTC CAGCTCTCGC GTGAGAAAAGC AAGCGGAGAG AAAAGAGGTC | 1440 |
| GCTTCCTACC TCTCCAAACG ATTCTTGGTT CCATAAGTTT AGGGATTGGC TATTATCTTG | 1500 |
| CCCTTACGGT AACCGATCCT CTTACAGCCC TAACAACCTT CTTCCTAGCT GTTTTGCTGG | 1560 |
| TTATCTTTGG TACTTATCTA TTGTTAATG CAGGGATTAC AGTCTTCCTA CAAATCTTAA | 1620 |
| AGAAAAACAA GAAATACTAT TACCAACCTA ATAACCTCAT ATCTGTTTCC AACTTGATTT | 1680 |
| TCCGTATGAA GAAAAATGCG GTTGGACTAG CAACCATCGC TATTTTGTC AATATGGTTT | 1740 |
| TGGTAACCAT GTCAGCAGCG ACAAGCATTT TCAATTCCGC AGAAAGCTTT AAAAAAGTTC | 1800 |
| TAAATCCTCA TGATTTTGGG GTTTCAGGGC AAAATGTTGA AAAAGAAGAT TTGGACAAAC | 1860 |
| TCTTGAGCCA GTTTGCAAGT GACAAAGGTT ATAGTGTC AAGAGAAAGAA GTACTTCGTT | 1920 |
| ACAGTAACTT TGGTATTGCA AATCAAGAAG GAACCAAGTT AACTATTTTT GAAAAAGGAC | 1980 |
| AAAACCGTGT CCAACCCACA ACAGTTTTC TGGTATTTGA CCAAAAAGAT TATGAAAATA | 2040 |
| TGACTGGTCA AAAACTGTCT CTATCAGGAA ATGAGGTCGG TCTCTTGCC AAAAAAGACG | 2100 |
| GACTGAAAGG ACAGAAAGCT CTAACCTCTA ATGATCATCA ATTTTCTGTC AAAGAAGAAT | 2160 |
| TTAATAAAGA TTTCATTGTG AACCATGTTC CAAATAAGTT TAATATCTTG ACTACTGATT | 2220 |

595

| | |
|--|------|
| ACAATTACCT TGTGTTCCCT GATTTACAAG CCTTTTGGGA TCAATTCCCA GATTCGGCTA | 2280 |
| TCTATAATCA GTTTTACGGT GGTATGAATG TAAATGTCAG TGAAGAAGAA CAACTCAAGG | 2340 |
| TCGCTGAGGA GTATGAAAAC TACCTCAATC AATTTAATGC TCAATTAGAC ACAGAAGGTA | 2400 |
| GCTATGTTTA TGGTAGCAAT CTAGCAGATG CTAGTTCTCA GATGAGTGCC CTCTTTGGTG | 2460 |
| GTGTCCTCTT TATCGGTATT TTCCTATCCA TTATCTTTAT GGTGCGAACT GTTCTGGTCA | 2520 |
| TCTACTACAA ACAAATTCTT GAAGGCTACG AAGACCGTGA ACGCTTTATT ATCTTGACAGA | 2580 |
| AAGTCGGTTC GGACCAAAAG CAAATCAAGC AAACCATCAA CAAACAGGTT TTAACGTGTT | 2640 |
| TCTTCCTTCC TTGCTCTTTT GCCTTCATAC ATCTCGCCTT TGCCTACCAT ATGCTTAGCC | 2700 |
| TGATTTTAAA AGTGATTGGT GTAGTGGATA CGACTATGAT GTTGATTGTG ACCTTGCTCA | 2760 |
| TCTGCGCTAT CTTCCTCATC GCCTATGTGC TGATTTTCAT GATTACTTCA AGAAGTTATC | 2820 |
| GCAAGATTGT GCAAATGTAA AAAAGATACC TCGACTTCAA AATCGAGGTA TTTCTTGAT | 2880 |
| TCTAAATGCT GAAAAGTTGT CCGAGCAGGA AGGTAAGTCC CATGGTCAAG AGACCAATAG | 2940 |
| CAAGGTTCCG AATCATAGCT GTTTTGGTTG GGGCTTTTCC AAGTCTAGCA CTTGTGTAAC | 3000 |
| CAGTGAGAAG AAGGGCCACA CCGACAATAA GGACGGTAGC AGGGATGCGG TAATCACTTG | 3060 |
| GAAAAATGGT CACTGACAGC ATTGGAGGCA AACTTCTAAG GAAAAAGGCA ACGAAGCTAG | 3120 |
| AAATGGCAGC GTGCCAAGGA TTGGTAAATT CTTCACTC AATCCCATAT TTTTCTCTA | 3180 |
| CCAGAGCCTT GAGTGGATTT TTAAGAAAGA TCTTATTGGT CAAGAGTTGG GCAGAAGTTT | 3240 |
| TGAATTCTCC ATTTTGGATA TAAGCAGCAT AGAGGGATTT TTTGGCTAGT TCCCTATCTT | 3300 |
| GGTCTAGCAA GAGTTTTTCT CGCGAAACGG CAGCTTCCTC GGTATCTTTT GGAGTTGAAA | 3360 |
| CGGATACATA TTCTCCACCA GCCATTGAAA AGGCACCAGC TAAGATAGCC GTAAAACCTG | 3420 |
| ATAAAAAGAT AATCCAGATA TTGGTCGTGG CACTGGCAAC TCCGATAACC ACACCAGCAA | 3480 |
| TGGAAATAAT TCCATCGTTA GCATCAAGAA CACCCGCACG CAGGATATTT AAACGACCTG | 3540 |
| CAAAATTTGA ATCAATTTTC TGATTTGTTT CTGACGCTAA ATTTCAAGTT CAAGTTAGCC | 3600 |
| ATCAAGAAGT CTTCTCTGGG TGACTTGTAG TCCAAGCATT TTTTAGGATA GTTGTTAATC | 3660 |
| CACTTTTCGA TGAATGCGAC TTCTTTGGGA GTCATTTTCT TGGTCCCTT AGGTAACCAT | 3720 |
| CTACGAATGA GCCTGTTGTG ATTCTCATTA GTTCCCCTTT CCCAAGAGGC ATAGGGATGT | 3780 |
| GCATAATAAA TGTGCTCCTC AGAAAATACA TTAGACAAGC GATTGAATTC CGTTCCATTA | 3840 |
| TCTGCCGTGA TGGAAAGAAT CTTGTGTGT TTTAAGATGA GTTTTAGAGC CTGATTGACC | 3900 |
| ACATCAGCAC TTTTATTTGG AATCAATCGG ATGATCTGAT GTCTACTTTT TCGATCCGTC | 3960 |

596

| | |
|--|------|
| AAGACAAGCA AGCAGTAGTT TTTCGCTCTC GTAAGTAGAA CTGTATCAAT CTCATAATGC | 4020 |
| CCATTCTCCA AGCGAAAATT GATAGCTTCA AGCCGCTGTT CGATGGATTG ACCAGCAGGT | 4080 |
| TTAAAGTTGG TGCTGGCCTG TTCTTTAAGC GCTTTTCCTT TTCTAGGGTA AAGCAGATCC | 4140 |
| TGTTTGCTTA ACCCCAATTT TCCATGATGA ATCCAATAGT AAATGGTTGA AATTCCCACG | 4200 |
| TTAACCCCTT TAGCCATCAC CATCATTTCA GCGGAAAATT TTGGTTATG ATAGTGGAGA | 4260 |
| ATCTTTTCCT TTAGTTCCTT GGTCAAGCTT GATTTCCTGA CCGAGCGCTT GCGATTGTTT | 4320 |
| TCATAAGACT GTTGAGCATA GTCGGCAGAA TAAACCTCTT TGAAGCGCCC TTTTCCAAGA | 4380 |
| CATTGTCGGA CTGTCCCACG CTTGATTTC A GTGTGGATAG TTTGAGGAAC TTTTCCAAGC | 4440 |
| AGAGAGGCAA TTTCTCTATT TGATTTCCTT TCTTTTTCCT ATCTTTCGAT TAAGCGACGG | 4500 |
| CTATCGATTG TCAAATGTTT GCCTTTTGTA GTATAATGGT TTTGCATCTC TGTGCCTTTC | 4560 |
| TTGTGTTTGT GGTGGAACAA CAAGTATAAC ACAGAGGTGT TTTCTTATGC CTACAAGAGC | 4620 |
| TATCGGCTAG TTGAACCATC TAATTTTTCG GAGGGCTGGG TGGCTAACTT' CATTATAGAA | 4680 |
| CTTTCATTTA CGAACATATA GTAAAATGAA ACAAGAACAG AACAAATCGA TCAGGACAGT | 4740 |
| AAAATCTATT TCTAACAATG TTTTAGAAGC AGAGGTGTAC TATTCTAGTT TCAATCTATT | 4800 |
| ATATTTTGTG TTTTATCAA AAAATACTTT ACAAGTTCTT AAAACATGA TATAGTAATA | 4860 |
| AAGCTTAGAA AATGAGATGA TGTTTCTAG CAAATATAAA CCCGAGTAAA AAATGCCTAC | 4920 |
| GGACAGGCAG GGTGAATGC CGAAGCGTGG TTGAAAAGCC ACATTATTGA TAGGGTTAAA | 4980 |
| AGCCTACTTT TATAAGTTGA TGTTAGGACA CTGTCTCTAA TTCATAAATT TTTAGTGTGG | 5040 |
| TGAAAGCACA CGTCATCTTG TGAAACGATC AATAAAGTAC GTAATATTG CTAAGTAGAGA | 5100 |
| GTTAGGAAAC ATCGGGAACA GACATACTCA ACAGAAACCA AAATAAACAC GTCAGAAGAT | 5160 |
| TGCAGAGCAG GTGAAAACCT GCTCTTTTTT CATGAGTCAA CCTTTAGTTC CTTAGTTTTT | 5220 |
| ATAAGTTCCT AAAAATATTG AAAGGAGTAT GTTTTGAAAG AGTTAGATCA AAACCAAGCC | 5280 |
| CCAATTTATG AGGCCTTGGT GAAGTTACGC AAGAAAAGGA TTGTCCCTT TGATGTTCCA | 5340 |
| GGTCACAAGC GTGGACGGG AAATCCAGAA CTGTGCGAAC TCTTAGGAGA AAAATGTGTA | 5400 |
| GGCATTGATG TCAATTCGAT GAAACCTTTG GATAATTAG GCCATCCTAT TTCGATTATT | 5460 |
| CGTGATGCAG AGGAGCTGGC TGCAGATGCT TTTGGAGCTA GCCATGCCTT TCTAATGATT | 5520 |
| GGTGAACAA CTTCATCGGT GCAGACTATG ATTCTGGCAA CCTGCAAGGC AGGAGATAAG | 5580 |
| ATTATTCTGC CACGAAATGT CCATAAATCT GCTATCAATG CGTTGGTTCT ATGTGGTGCC | 5640 |
| ATTCCCCTCT ATATCGAGAT GAGTGTAGAT CCTAAGATTG GTATCGCTTT AGGTCTTGAA | 5700 |
| AATGACCGAG TAGCACAGGC CATAAAGGAC CATCCAGATG CTAAGGCTAT CCTAATCAAC | 5760 |

597

| | |
|--|------|
| AATCCTACTT ACTACGGCAT CTGTTTCAGAC CTAAAGGGGT TGACAGAAAT GGCTCATGAA | 5820 |
| GCTGGCATGA TGGTTTTAGT AGATGAAGCC CACGGAGCGC ATTTGCATTT CACTGATAAA | 5880 |
| CTTCCAATTT CTGCTATGGA TGCAGGGGCT GATATGGCAG CAGTTTCCAT GCATAAGTCT | 5940 |
| GGTGGGAGTT TGACCCAAAG CTCCATTTTA CTTATCGGGG AGCAGATGAA TTCTGAATAC | 6000 |
| GTTCGTCAGA TAATTAACCT GACCCAGTCT ACATCTGCCT CTTACTTGTT GATGGCTAGT | 6060 |
| TTGGATATTT CACGTCGCAA CTGCGCCCTT CGTGGTAAAG AGTCGTTTGA GAAAGTCATT | 6120 |
| GAGCTATCTG AGTATGCCCC CCGTAAATC AATGCTATCG GTGGCTACTA TGCCTACTCA | 6180 |
| AAAGAGTTAA TAGACGGTGT TTCGGTTTGC GATTTTGACG TAACTAAGCT GTCAGTTTAC | 6240 |
| ACTCAGGGTA TTGGCTTAAC AGGTATCGAG GTTTATGACC TCTTGCGAGA CGAATACGAC | 6300 |
| ATTCAGATCG AGTTTGGTGA TATCGGCAAT ATCTTGGCCT ATATTTCCAT CGGCGACCGC | 6360 |
| ATCCAAGACA TCGAGCGCTT GGTGGTGCT CTGGCTGATA TTAAGAGACT CTATTCAAGA | 6420 |
| GATGGAAAAG ATTTGATAGC AGGAGAATAT ATTCAGCCCG AGTTAGTGCT GTCTCCGCAA | 6480 |
| GAAGCCTTCT ATTCAGAAAAG AAAAAGTTTA ACTTTGGATG ATTCTGTTGG ACAGGTCTGT | 6540 |
| GGAGAATTTG TTATGTGTTA CCCTCCAGGT ATTCCTATCT TGGCTCCTGG TGAACGCATT | 6600 |
| ACACGAGAAA TTGTGCACTA TATCCAATTC GCCAAGGAAC GTGGTTGCTC CCTCCAAGGG | 6660 |
| ACGGAAGATC CAGAGGTCAA TCATATCAAC GTTATTAAGA GAAAGACAAA CTATAAGAAA | 6720 |
| AGTCAATAGT TTTATCTAAA CTATTTCTTA TTTCATTTG ATGATTTGGC GATGATTTTA | 6780 |
| GAGCACGGCA AAAAGCCCTT GAATTAGAAG CGGTCAATCG CTTAATTTCT ATCAGCTTAT | 6840 |
| CAAATCCTGC CTCAAGCCTT TTCTGAGGAT TAGGGTAGCG TGTCAAGAGT TGGTAGGTAT | 6900 |
| ATTCCTGAATG CTTTCCAACG ATTTTATCCA ACTCAGGAAA GATGATATCA AGACAACGAG | 6960 |
| TGTATTGTAC TTTCCAATCA GACTGTTTTT TCTTGAGACG ATGAATATGT CTAGCCAGTA | 7020 |
| TTTTTAGTTC TACTTGCCGA TTATCGTGTG GAAATGTTC ACGATTGGGG TCAGAAAGAA | 7080 |
| GTTTAAGAGC GATGCCATGA GCGTCTTTCT TATCCGTTTT AGTTTTCGA AGTGATAATG | 7140 |
| ATTTGGCAAA TTTCTTGATG AGCAAAGGAT TGTAGGTGTA AACTTTATAT CCTTGTTTAT | 7200 |
| GCAGGAAGTT CAGTAGATTA AAGGCATAAT GTCCGGTATT TTCAAGAGCG ATGAGACAGT | 7260 |
| CTTGGTTGAG CTGTCGAAGA GACAGATCTA AGAGTTCAAA ACCAGCTTTA TTATTTGAAA | 7320 |
| AAGTGAGTGG TTTAAGAACA GTTTTTCCTG GAACATTCAA GGCTGTAACA TCGTGTTTAT | 7380 |
| TTTLAGCGAC ATCAATGCCC ACATAAAGCA TGGGAGTATC TCCAGATATA GTATTTCAAG | 7440 |
| TCTACTGGGT TATCCACGAA CTTTTGCTT TGTACCTTA GACGAGATAA AACGTCTATG | 7500 |

598

| | |
|--|------|
| CGTTATCAAA CTCATTACCA ATTGAAACAA AAAACTGTGG TTAGAGCCTT TCGGAAATCG | 7560 |
| TCAAGCGATT GGAGGAAATG AACTAATCCA CAGTGGCTTA TTCCAAGTAT ACCACTTGGG | 7620 |
| CTTTGGCAGT AGCTAACTGC GCTAAATATA ATATAAGGAG AAATAGATGG ATTTATGGTT | 7680 |
| TTCTGAAGTT CATACTCCAG ATGTCAAAT GTCTCTGAGA ACAGCCAAGC AACTTTACGC | 7740 |
| TGGAAAAAGT GAATGGCAGG ATATCGAAGT CTGCGATACG CCAGCTTTTG GGAAATACT | 7800 |
| GATTTTAAAT GGCCATGTCT TGTTCCTCAGA TCGGATGAT TTCGTCTACA ATGAAATGAC | 7860 |
| CGTTCACGTT CCCATGGCTG TCCACCCAAA TCCAAAGAAA GTATTGGTTA TTGGGGGTGG | 7920 |
| TGACGGCGGT GTTGCCCAAG TATTAACCCT CTATCCTGAA CTGGAGCAAA TTGATATTGT | 7980 |
| GGAACCGGAT GAGATGTTGG TCGAGGCTCG TCGTGAGTAT TTCCCAGACT TTGCTGCAGG | 8040 |
| GCTAGATGAT CCTCGTGTTA CCATTTACTA CCAAAATGGG CTACGCTTTT TCGGAAACTG | 8100 |
| CGAAGATGAT TACGATATTA TCATCAACGA TCGACAGAT CCATTTGGCC ATACGGAAGG | 8160 |
| ACTCTTTACC AAGGAATTCT ACGGCAATAG TTATCGAGCT CTGAAGGAAG ACGGCATCAT | 8220 |
| GATTTACCAG CATGGGAGTC CCTTCTTTGA CGAGGATGAG TCGGCCTGCC GAAGCATGCA | 8280 |
| CCGCAAGGTC AATCAAGCCT TTCCAATCAG TCGGGTTTAT CAGGCCATA TTCCAAC TAG | 8340 |
| CCCAGCTGGC TATTGGTTGT TTGGATTGTC ATCGAAAAA TACCACCCTG TCAAAGATTT | 8400 |
| TGACAAGGAA GGCTGGAAAA AACGCCAGCT TTTACAGAA TACTACACTG CAACTTACA | 8460 |
| CGTGGGAGCC TTTATGTTGC CCAAGTATGT TGAGGACATT TTAGAAGAAG AGGAAGGAAA | 8520 |
| AAAATGAGTC GTTTACTAGT TATTGGTTGT GGGGCGTTG CCCAAGTTGC TATTTCAAAG | 8580 |
| ATTTGTCAAG ATAGCGAAAC ATTTACAGAG ATTATGATTG CTAGCCGTAC CAAGTCAAAA | 8640 |
| TGCGATGACT TGAAAGCGAA GCTAGAAGGC AAAACAAGTA CTAAAATTGA AACTGCAGCA | 8700 |
| CTTGATGCTG ACAAGGTTGA AGAAGTGATT GCCCTGATTG AAAGCTACAA ACCAGAAGCT | 8760 |
| GTTTTGAATG TAGCTCTGCC TTATCAAGAT TTAACCATTA TGGATGCTTG TTTGGCAACA | 8820 |
| GGTGTTCACT ATATCGATAC AGCCAAC TAC GAAGCAGAAG ACACAGAAGA CCCTGAGTGG | 8880 |
| CGTGCTATCT ACGAAAAACG TTGTAAGGAA CTTGGTTTTA CAGCCTACTT TGACTACTCA | 8940 |
| TGGCAGTGGG CTTATCAAGA GAAATTCAAA GAAGCAGGCT TGACTGCTCT TCTTGGTTCT | 9000 |
| GGTTTTGACC CAGGTGTAAC TAGTGCTTTT TCAGCTTATG CCCTCAAACA CTATTTTGAT | 9060 |
| GAAATCCATT ATATCGACAT TTTAGACTGT AATGGCGGTG ACCACGGTTA TCCATTTGCA | 9120 |
| ACCAACTTTA ATCCAGAAAT TAATCTCCGT GAGGTTTCTG CGCCAGGTTT TACTGGGAA | 9180 |
| GATGGGAAAT GGGTCGAAGT CGAAGCTATG TCTATCAAGC GTGAGTATGA TTTCCCTCAA | 9240 |
| GTTGGACAAA AAGATATGTA TCTCCTTCAC CATGAAGAAA TCGAATCATT GGCCAAGAAC | 9300 |

599

| | |
|--|-------|
| ATTCAGGTG TCAAACGCAT TCGTTTCTTT ATGACTTTTG GTCAATCTTA CTTGACGCAC | 9360 |
| ATGAAATGTC TTGAAATGT TGGACTCCTT CGTACGGATA CCATTAACCT TAACGGCCAA | 9420 |
| GAAATTGTTC CAATTCAATT TTTGAAAGCC TTGCTTCCAG ATCCTGCCAG TCTTGGGCCA | 9480 |
| CGTACAGTCG GAAAAACCAA TATTGGATGT ATCTTTACAG GTGTCAAAGA CGGTGTCAAA | 9540 |
| AAGACTATCT ATATCTACAA TGTCTGCGAC CATCAGGAAT GTTACGCAGA GGTGGTTCCG | 9600 |
| CAAGCTATTT CTTATACGAC AGGAGTTCCA GCCATGATTG GGACAAAATT AGTCATGAAC | 9660 |
| GGAACCTGGA AACAAGCTGG AGTGTATAAC CTTGAGGAGT TAGATCCAGA TCCATTTCATG | 9720 |
| GAAGCTTTGA ATGAGTATGG TTTGCCATGG GTTGTGGTTG AAAATCCACA AATGGTGGAC | 9780 |
| TAATGAAGTT AGAACAAGTA CCAACACCAG CCTATGTTAT TGACTTGGCC AAGTTAGAAG | 9840 |
| CTAATTGCCG CATTCTACAA TATGTACAAG AAGAGCCCG TTGCAAGGTC TTGCTTGCCC | 9900 |
| AGAAGGCATA TTCCCTCTAC AAAACTTATC CCTTGATTAG CCAGTATCTA TCAGGTACGA | 9960 |
| CAGCTAGTGG ACTCTATGAG GCCAAATTGG CAAGGAAGA ATTTCTGGT GAAGTCCATG | 10020 |
| TATTTGCGCC TGCTTTCAAG GATGCAGACT TGGAGGAATT GCTAGAGATA ATGGACCATA | 10080 |
| TAGTCTTTAA CTCAGAGAGA CAGTTGCGTA AACACGGTCC GCGTTGTCTGA GAGGCTGGTG | 10140 |
| TCAGTGTGG TTTGCGCCTC AACCTCAGT GTTCAACTCA AGGcAGATCA CGCGCTCTAT | 10200 |
| GACCTTTGTG CACCAGGTTT TCGCTTTGGA GTTACTATAG ACAAGATTCC GAGTGATTG | 10260 |
| CTAGATTGG TTGACGGACT TCATTTTCAT ACCCTTTGCG AGCAGGGAGC AGATGATTTA | 10320 |
| CAAACAACCT TGAAAGCAGT AGAAGAACAG TTTGGTCCCT ACTTACATGA GGTAAATGG | 10380 |
| CTCAATATGG GTGGTGGTCA TCATATTACA AGAGAAGTT ACGATGTGGA TTTGCTGATT | 10440 |
| TCAGAAATCA AGCGTATCCG AAAAATTAC AATCTTGAAA TCTATATCGA GCCTGGTGAA | 10500 |
| GCCATTGCGC TTAATGCGGG TTATTTAGCA ACTGAGGTAT TAGATATTGT AGAAAACGGT | 10560 |
| ATGGAAATCT TGGTTTTAGA CGCCTCTGCG ACCTGCCATA TGCCGTATGT ACTTGAGATG | 10620 |
| CCCTATCGTC CACCTTTGAG AAATGGCTTT GAGTCACAGG AAAAAGCCCA TACCTACAGA | 10680 |
| CTTTCTTCTA ATACCTGTCT GACGGCGAT GTGATTGGTG ATTATAGTTT TGAAAATCCA | 10740 |
| GTCCAAATCG GAGACAGACT TTATTTTCAA GACATGGCCA TTTATCTTT TGTCAAAAAT | 10800 |
| AATACCTTTA ATGGTATTGG ATTGCCAAAGT CTCTATCTCA TGGACGAACA GGGAGACTGT | 10860 |
| AGCTTACTCA AAGCTTTTGG CTATCAAGAC TTTAAAGGA GATTATCATG ATGGACAGTC | 10920 |
| CAAAAAAATT AGGCTATCAC ATGCCAGCAG AGTACGAACC CCATCATGGT ACCCTCATGA | 10980 |
| TATGCGCGAC TCGACCAGGA TCATGGCCTT TTCAAGGAAA GGCTGCTAAA AGAGCATTTA | 11040 |

600

| | | | | | | |
|------------|------------|------------|------------|-------------|------------|-------|
| CTCAGATTAT | CGAGACCATA | GCAGAAGGGG | AAAGAGTCTA | TCTTTTGGTG | GAGCAGGCCT | 11100 |
| ATCTATCTGA | AGCCCAATCC | TATCTTGGAG | ACAAGGTTGT | TTATTTAGAC | ATTCCCACCA | 11160 |
| ATGATGCCTG | GGCGCGTGAT | ACTGGCCCAA | CCATTCTCGT | CAATGATAAA | GGTAAGAAAT | 11220 |
| TAGCCGTGGA | TTGGGCCTTC | AATGCTTGGG | GAGGCACCTA | TGATGGTCTT | TATCAAGATT | 11280 |
| ATGAAGAGGA | TGACCAAGTA | GCCAGTCGTT | TTGCTGAGGC | CTTGAAAGG | CCTGTCTATG | 11340 |
| ATGCTAAACC | TTTTGTACTG | GAAGGAGGCG | CAATCCATAG | CGATGGTCAA | GGAACATTC | 11400 |
| TCGTAAGTGA | AAGTTGCTTG | CTTAGTCCTG | GTCGCAATCC | TAACCTTGACT | AAAGAGGAGA | 11460 |
| TTGAAACAC | ATTATTAGAA | AGTCTTGGTG | CTGAAAAAGT | TATTTGGCTT | CCTTATGGTA | 11520 |
| TTTATCAGGA | TGAAACCAAT | GAACACGTCG | ATAATGTTGC | TGCCTTTGTT | GGTCTGCTG | 11580 |
| AGCTTGTTTT | GGCTTGGACA | GATGACGAAA | ATGATCCCCA | GTATGCCATG | TCAAAAGCAG | 11640 |
| ATCTCGAACT | CTTAGAACAG | GAAACAGATG | CAAAAGGTTG | TCACTTCACC | ATTCATAAAT | 11700 |
| TGCCTATCCC | TGCAGTTCGA | CAAGTTGTGA | CAGAAGAAGA | TTTGCCAGGC | TACATCTATG | 11760 |
| AAGAAGGAGA | AGAAAAGCGA | TACGCAGGTG | AACGACTAGC | AGCTTCCTAC | GTAAACTTTT | 11820 |
| ATATCGCCAA | CAAGGCTGTC | TTGGTTCCAC | AGTTTGAGGA | TGTAAACGAC | CAAGTGGCCT | 11880 |
| TAGATATCCT | CAGCAAGTGT | TTCCCAGACC | GTAAAGTTGT | CGGAATACCA | GCCAGAGATA | 11940 |
| TTCTCTTAGG | TGGTGGCAAT | ATCCACTGTA | TCACCCAACA | AATTCCAGAA | TAGGAGAAAA | 12000 |
| AGATGAGAAA | TGTAAGAGTT | GCAACCATTC | AGATGCAATG | CGCTAAGGAT | GTGGCAACAA | 12060 |
| ATATCCAAAC | CGCAGAGCGT | TTAGTACGTC | AGGCTGCTGA | GCAAGGAGCC | CAAATTATTC | 12120 |
| TCTTGCCCGA | GTTGTTTGAA | CATCCCTATT | TCTGTGAGGA | ACGTCAGTAT | GACTACTACC | 12180 |
| AGTATGCCCA | ATCTGTAGCG | GAAAATACTG | CCATTCAGCA | TTTTAAGGTG | ATTGCTAAGG | 12240 |
| AACTACAAGT | TGTTTTACCA | ATCAGTTTCT | ATGAAAAAGA | TGGTAATGTC | TTGTATAACT | 12300 |
| CTATTGCCGT | CATTGATGCA | GATGGGGAAG | TGCTGGGCGT | TTATCGAAAG | ACCCATATAC | 12360 |
| CAGATGACCA | TTATTATCAA | GAAAAATTCT | ATTTACAGCC | TGGTAACACT | GGTTTCAAGG | 12420 |
| TCTGGAATAC | TCGCTATGCT | AAGATTGGTA | TCGGTATCTG | TTGGGATCAA | TGGTTCCCTG | 12480 |
| AAACAGCGCG | CTGTCTTGCA | TTGAATGGTG | CTGAATTGCT | CTTTTATCCT | ACAGCTATCG | 12540 |
| GTTCAGAGCC | AATTTTGGAT | ACAGATAGTT | GTGGTCACTG | GCAACGTACT | ATGCAAGGGC | 12600 |
| ACGCAGCAGC | GAATATTGTT | CCAGTCATCG | CAGCCAATCG | TTATGGTTTA | GAGGAGGTTA | 12660 |
| CTCCTAGTGA | GGAAAATGGC | GGACAGAGCT | CCAGTCTTGA | CTTCTACGGT | TCCTCCTTTA | 12720 |
| TGACGGATGA | AACAGGAGCT | ATTCTAGAAC | GAGCTGAAAG | ACAAGAAGAA | GCTGTTCTGT | 12780 |
| TAGCTACTTA | TGACCTAGAC | AAGGGAGCAA | GTGAACGCCT | AAACTGGGGC | TTGTTTCGAG | 12840 |

601

| | |
|--|-------|
| ATAGAAGACC AGAAATGTAT AGACAAATTA CAGATTAGTG TGGGAGAAAT GAGAGATTCA | 12900 |
| TTCTGCTAGA CTAACCTCTT ATTAGTAACT ATAAGATACT ATGGCATCTA GTAAATCGAT | 12960 |
| TTTTATGATT CGCTATTCTT GTCTATTGAT TAGTCCGTAT TTTAAAAATAT TAGCAAAAAA | 13020 |
| GCAAATAGCA GTAACCTCTG TCTATTGCT TTTCTTTTTT ATAGAATATA TTTCTCAATA | 13080 |
| GCACGCGCAA CGCCGTCTTC TTCGTTGCTT GAGGTAACGG CATCCGCAAG AGATTTGATA | 13140 |
| TAATCGCTGG CATTTCCCAT TGCAATCCCA AGCCCTGCAA ACTGGAGCAT TTCGATATCG | 13200 |
| TTATTAGCAT CGCCCATGGC CATAATCTCT GAGGAATCAA TCTTCAAAAT CTCAGCTAGT | 13260 |
| CGTGAAAGAG CAGTAGCCTT TGTCGTTCCA AGCGGCATTG CTTCAATAAT GACAGGCTGC | 13320 |
| GAACGAACTC CACTGAATCG TTGGCAAAGC TCTTCAGCAA AACGCTGCTC AAAATCGTCT | 13380 |
| GTTTGTCTT TTGTTCTTAA ACACATACCT TGGAAATCC GGAACCTTCC ACTAGTCGCT | 13440 |
| TCTTCAAGAG AAATTTCACT CAGGTCTGAA AATACTAGTT TAGCATCATT TTCAATAACT | 13500 |
| TGATTGGGCT TGTCACCGAG AACAAAATAA TGTGACTCGT CAAAAAGTGT CAACTGAACA | 13560 |
| TCACTCTTTT CAGCAAGGTC ATAGAGGTAT TCGATGTCAG CTGGACTCAG TTCTTTCCAG | 13620 |
| TCAACTAGAC TCCAATCACT GGTCTGGTGA GTTGAACAAC CGTTGTTAAC AATAATATAT | 13680 |
| TCGTTCTGGA GGTCAAGCTC CAGTTTTTTG TAGTAGGGA GGACACCGAA AAGGGGGCGA | 13740 |
| CCCGTACAGA GAACCACTT GACACCTTTT TCAATGGCTT TGTGAATAGC AGTAATGTGT | 13800 |
| GCTTGTGGGA TTTCTTGGC TTCATTGAGG AGGGTGCCGT CCATATCCAA GGCTAGTAGT | 13860 |
| TTAATCATAG GTCTTCCTCT TTATCTTTGC TATTATTATA GCATATTTG GAGAAGAAAT | 13920 |
| TGATAGAAAG CTTGAGACTA ATTGATTTTA TAGTTTAAGA TGTTTTGATG ACAATTCATG | 13980 |
| ATTTGAAGAG GATATTTGCG AAAGATATGC TATACTATGT TTGTCAATGT TGCAACTAGA | 14040 |
| CAAATTAAAA AACCAACTTA ATATAATAGT TTTTTGTAA GTAGGTATGA GTAGCAGATT | 14100 |
| ACTCAACTAA TCTGAAGAAT AATGGAGGAA ATATATCATG ATTTTAATGA CAAAAATAT | 14160 |
| AAATCTAACA AATGAAGAAT TAGAGCTGAT ACAAGGTGGA GCAGATCCAT ATGGTAAAAA | 14220 |
| TCCTAATGGT AGGTACGATT GGGAAATAGA ACCAGTATTA ACTCTGCTGG TTCATGGATT | 14280 |
| TTGTCCCAGA GGCACCTATG ATTCAGGATA TATTGGAGGA GGTAAATCATC TTTGCAAAGG | 14340 |
| AAGTGCTGCG AGATTTTAAG TAAAAATTAT TAGGAATATG AAGAAACAAG GGGAGAAAAC | 14400 |
| AGAGGATTTA ATATGAAAAA ACGAGCTATT CAAATTTTAC TAGCATTTGC CTTAATTTTT | 14460 |
| TACAAATCAA CTTGGTTTTG GAGGCTTTTC AATTATCTCG CAAAGCCCTA TCTACCAGCA | 14520 |
| AGTCGTGAAT TTTTTCAGAT TCTGCTTTTG ATGGAGAGCG GAGTTCTTTT CTTAGCGGTC | 14580 |

602

| | |
|---|-------|
| ATCTATCTAC TGGTTTTTCG AGGAAAGAAA ATTTTTCATT TCAAGTGGCA GCTGAGGTAC | 14640 |
| TTCATCTACC TTTTACTGGG CTACATCATT TCATATATGT CTGACTTCCT CTMTTCGTAT | 14700 |
| TTCATATCCC TGTCTTCAA TCAGATTCTT TTGAATGAAA CGGTAGAAAT GATGGGGAGA | 14760 |
| CAGGAGTTCC CTTATGTCTT GCTCATCGTT TGCTTCATCG CCCCTATTGC TGAGGAATTG | 14820 |
| ATTTATCGAG GtGTGCTTAT GACAACCTGT TGCAAAAAC CACCTTGGTA CG | 14872 |

(2) INFORMATION FOR SEQ ID NO: 73:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10223 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

| | |
|---|------|
| CGTGCTATCG GTCTCAAAAC CAATCTGGTC GCTATGGTCA AATCCAGTTG GAAAATCCAT | 60 |
| TCTTCTTGGA GCCATCTGCT GGATTGCCAT CATCCTCACC ACTCTTGGTA TGCAGACCCT | 120 |
| TATCGGCATT TTCTAATACT CTTGCAAAAT CTCTTCAAAC CACGTCAACG TCGCCTTGCC | 180 |
| GTAGGTATAT GTTACTGACT TCGTCAGTTC TATCTGCAAC CTCAAAACGG TGTTTGAGCT | 240 |
| GACTTCGTCA GTTCTATCTG CAACCTCAA ACGGTGT TTT GAGCTGACTT CGTCAGTCGT | 300 |
| ATCTACAACC TCAAAACAGT GTTTTGAGCT GACTTCGTCA GTTCTATCTG CAACCTCAA | 360 |
| ACAGTGT TTT GAGCAGCCCG TGGCTAGTTT CCTAGTTTGC TCTTTGATT TCATTGAGTA | 420 |
| TAACACAAAA GGTAGCCCAT CAGCTACCTT TTTCTTATGC TTCTCAATC AAGCGAGTAT | 480 |
| GTTCTCTCTT GATACAGCGA TTCATCACGA TATCATCACA TCCACCATCA CGCAAAATCT | 540 |
| CTTTCGCTTC TAAACTTTCA AGTCCTAGCT GTGCCCCAAA AATCTTGGCA TCAGCTTTGA | 600 |
| GAAAATCACG CGCCACATCG GGCAGAAATT CACTGCGACG ATAAACATTG ACAATATCTA | 660 |
| CAGGAAAAGG AATTTCAGCG AGGCTAGCAT AAGCCTTTTC ACCCAAGATT TCGCCACCTG | 720 |
| CCGCCTTGGG ATTGACTGGG ATGATTTTAT AGCCCCGAGC CTGCATTTCC TTTGTTACTC | 780 |
| GATTGCTGGT TGTTCCTTCA CGGTCAGACA AACCCACCAC AGCAAGGGTT TTA CTCTGTTG | 840 |
| CGAGATACTG ACGAATCACG CCATCACTTG GATTGATAAA TTCTTGACTC ATAGAAATCC | 900 |
| TCCTTTTTC TCA GTATAGC ACATTTTGAA AAGGTTTGCA GAATTATACT ACAAAAAGG | 960 |
| AGGACTAGCC CCTTTTAT TTAGCCTCGT ACCAGGTTGC CCCTTCATTC TCATCTGCGA | 1020 |
| TAAGAGGAAC ACTGAGTTGA ATGGCTTCTT CCATGGTTTG TTTCACCAAT TTTTTCATCT | 1080 |
| CTACCAATTC AGATTTAGGC ACTTCAAGGA CGATTTTCATC GTGCACTTGT AACAGCATCT | 1140 |

603

| | |
|--|------|
| TAGTCTGATA ACCACCTGCA ACCAAGGCTT TATCCAGCTG AATCATGGCA ATCTTGAGAA | 1200 |
| TATCTGCTGC CGAACCCTGG ATAGGTGAGT TGATAGCAGT TCGCTCCGCA AAACCACGAA | 1260 |
| TATTGAAGTT GCGCGAATTG ATATCTGGCA ACTCACGGCG ACGCTTAAAG AGGGTCTCTA | 1320 |
| CATAGCCCTT ATCACGCGCC TCCCGCACCA CTTTCATCCAT GTAGTTTTTA ATACCTGGAA | 1380 |
| AACGTTCAAA GTAGGTATCA ATGTAGGCTT TGGCTTCCTT ACGACTAATT CCCAAATTAT | 1440 |
| TAGACAAGCC AAAGTCTGAA ATCCCATAAA CCACTCCAAA GTTAACTGCC TTGGCATTGC | 1500 |
| GACGGTCGTT TGCAGTCACA TCATCAGGAC GCTCAATGCC AAAGACCCGC ATGGCTGTGC | 1560 |
| AAGTATGGAT ATCTGCCCCC TCTTGGAAGG CCTTAATCAA GTGCTCATCC TTAGAAATAT | 1620 |
| GCGCCAAAAC GCGCAATTCA ATCTGTGAAT AGTCAGAGCT GAGTAGCACA CTATCCTCCC | 1680 |
| ACTCTGGCAC AAAAGCCTTC CGAATCAAGC GCCCCTGTTC CAATCGGGCA GGAATATTTT | 1740 |
| GCAAGTTTGG ATCCACACTA GACAAACGCC CGGTCTGGGT CAAATCCTGC ACATAGCGAG | 1800 |
| TATGAATCTT TCCATCAGCC AAAATCCAGT CCTGCAAGCC AATTACATAA GTAGATTGAA | 1860 |
| TCTTAGCAAT TTGACGGTAA TCCAGGATTT TCTTAACAAT CGGAGCAATA GGAGCGAGAC | 1920 |
| GCTCTAAAAC ATCCACTGCT GTCGAATAAC CTGTCTTGGT TTTCTTAGTG TATTCTAGAG | 1980 |
| GAAGTCCCAA TTTCTCAAAG AGAAGCACGC CCAACTGCTT AGGCGAGTTG ACATTAAACT | 2040 |
| CCTCACCAGC CAGCTCGTAA ATCTCTTGAG TCAGTTTTTC AATGACAAGC TCATTTTCAG | 2100 |
| CCTGCATCTC AAGCAAGGTC TCTTTCTTGA CCATAATCCC AGCAATTTCC ATCTTGGCAA | 2160 |
| GGACAAAAGC CAGAGGTTGC TCCATATCAT AAAGAAGCTC TAATTGCCCA TTTTCGCTGA | 2220 |
| GTTTTTCAAG TAAAATAGGC TCTGTTTCTA CAAAACAGC AAGTTTACAA GCTAAGTGTT | 2280 |
| CCAAGAATTT CTCACGTTCA GGAATGGCCT TTTTAACACC CTTACCGTAG AAAGTTTCAT | 2340 |
| CATCAACCAA GTAAGTCTGA CCATAAAGAC TAGCGATGGT CGCAATTTCA TTGTCCTCCA | 2400 |
| CAGTCGAAAG GAGGTATTTA GCCAAACGGA TGTCAAAAGC AGGCGCCTGC AAATCCACAC | 2460 |
| CAAAACGTTG CAAAAGAACT TTAACCTTCT TAAAGTCATA AACTCTCAGA GATGTTTTTT | 2520 |
| CTAAGAAATC CTTGAAAATC GGTCTTGCA ACAGCTCAAG CTTGTCTGTG GCATAGAGCT | 2580 |
| TATCCCCACA AGACCAGACA AATCCAACCA AATTATCCGT ATGGTAATTC TCACCAAAAA | 2640 |
| GCTCAAAGTG GAAGATAGAC TCTTCACTCA GCATATCTTG ACTGATTTGG TCAACAATAG | 2700 |
| TAAAATCCAA ACTCTCAGAC ACATCAGCTG ACGACACATT TAAAGCCTGC TTTAGCTGTT | 2760 |
| TGAAGCCCAT CTCATCGTAG AATTTCCCAA GATTTTCAAC ATCTGGACCA CTATAGACCA | 2820 |
| AGTCCTCTAA ACCAATCGCA ATCGGTGCCT TGGTATCAAT GGTGCTAGT GTTTTAGACA | 2880 |

604

| | | | |
|------------------------|-----------------------|------------------------|------|
| AAAAGGCCTG TTCCTTGTC | TTGATGAGAT TTTCTTCAT | CTTAGAAGTC TTCATTCCAT | 2940 |
| CAATATTTTC ATAAATCCCC | TCAAGCGAAC CATGCTCCAG | CAAGAGCTTA ATACCCGTCT | 3000 |
| TTTCACCGAC TTTGGTCACC | CCAGGGATAT TATCCGACTT | ATCACCCATG AGCGCCTTGA | 3060 |
| GATCGATAAA CTGAGCTGGT | GTGAGGCCCA TTTCTTCCAT | GAGGTAATCT GCGTAAAGG | 3120 |
| CCTCAAACCT AGCCACACCT | TTCTTGGAAT TTTCAACCAC | CGTATGCTCA TCCGTCAGCT | 3180 |
| GAATCAAATC CTTGTCCCCA | CTGACAATAG TAATATCAAA | ACCATCCTGC TCTGCTAGCT | 3240 |
| TATCCAGCGT CCCAATGATG | TCATCCGCCT CATACTGAGC | CAGATCATAG TGACGAATCC | 3300 |
| CCATATGATC CAGCAACTCA | CGAATGAAAG GAAATTGCTC | ACGAAACTCA TCAGGAGTCT | 3360 |
| TGGCCCGACC ACCCTTATAG | TCCGCATACA TCTCTGTCCG | GAAGGTCGTC TTTCCCGCAT | 3420 |
| CAAAAGCCAC CAAAATATGA | CTCGGCTCAA CCCGCTCCAA | TAAATGACTC AACATCAACT | 3480 |
| GAAAACCATA AATCGCATTG | GTATGCAAAC CAGCCACATT | CTTAAAACGG TCCAACCTGCT | 3540 |
| GATACAGCGC AAAAAACGCC | CGAAAAGCTA CAGAAGACCC | ATCAATCAAT AATAATTTTT | 3600 |
| TCTTATCCAT ACACCCATTA | TAAAGGAAAG AATCAAAAAA | TACCATTGGG AAGAGCTAGA | 3660 |
| GCAAGTATTT TTCAACTTTT | TTCCGAATAA ATAGATAGAG | CCAGAGAATT TAGTAAACCT | 3720 |
| AGATTTAAAA ATGTGCTATA | ATATAGTATA TTGAATCTAT | AATAGTACAC CTTGACTGCT | 3780 |
| AAAATATTTT TATAAATTAA | TTTGACTTTC CTGATAGAGT | TATTCACATC TTATTTCAAC | 3840 |
| TCACTATAGA AGGAGGAATA | GGAGGATTCT CAGACATCCG | GGCATCAGCC CAACTAATGA | 3900 |
| TTTGATTGCT AAGAAAATAT | TCAGCAATCC AGAAATCACT | TGTCAATTTA TTCGCGATAT | 3960 |
| GCTGGACTTG CCAGCAAAAA | ATGTGACCAT TTTGGAGGGA | AGCGATATTC ACGTATTACT | 4020 |
| CTCCATGCCT TACTCGGTGC | AGGATTTTTA TACCAGTATA | GACGTCTTGG CGGAGTTGGA | 4080 |
| TAACGGTACT CAAGTAATTA | TTGAGATTCA AGTCCATCAT | CAGAATTTTT TCATCAATCA | 4140 |
| CTTGTTGGGCT TACCTGTGCA | GTCAGGTTAA TCAAAATCTT | GAAAAAATTC GTCAGCGAGA | 4200 |
| AGGTGATACT CACTAGAGCT | ACAAACACAT CGCTCCTGTT | TACGCCATTG CTATCGTGGA | 4260 |
| TAGTAATTAT TTCTCAGATG | ACCTGGCTTT TCATAGCTTT | AGTATGCGCG AAGACACAAC | 4320 |
| AGGTGAGGTA TTGGCGATTA | CCAACAATGG ACAGGAAAAC | CATCTGGTTA AGATGGCATT | 4380 |
| CTTGGAATTA AAAAATACAG | AGAAACCAGC AAAGACAAGG | TTGCAAGCC ATGGTTGGAG | 4440 |
| TTTTTCGGCA ACAAGCCCTT | TACCCAGCAA CCGCAACGAG | CCATTACCCA AGCAAAATCA | 4500 |
| CTGCTGGACT ACAAGAGCTG | GTCCGAGGAG GACAGGAAAA | TGTTTAGTCA ACTACATATG | 4560 |
| CGAGAAGAAC AAGTCTTGT | AGCACAGGAC TATGCCTTGG | AAACTGCTAG GGCTGAAGGC | 4620 |
| CTTGAACAAG GACTAGAGCG | TGGGAAAGTT GAAGGAAGGG | CAGAAAGGAA ACTTTTTGCC | 4680 |

605

| | |
|--|------|
| TTCCTAGACA TAGTACGCCA AGGTCCTCTG ACTTCTGAGG TTGCCAGCCA GCAATTAGGT | 4740 |
| ATGTCACTAT CTGAATTGGA GGCACGTGTG TAAAATGGCT CCATAATATC CATAGTGGGT | 4800 |
| AAATCCCCCTA TGGATATTAT GGAGCCTATT TTGTGTAGAA AAAAAGTCCC ATATGACCTA | 4860 |
| TAATGAAAAG CGACAAAACA ACTCATTAGA AAGAATCATA TGGAACAATT ACATTTTATC | 4920 |
| ACAAAATTAC TAGACATTAA AGACCCTAAT GTCCAGATTT TAAACATCAT CAATAAGGAT | 4980 |
| ACACACAAGG AAATCATCGC CAAACTGGAC TACGACGCCC CATCTTGCCC TGAGTGCCTA | 5040 |
| AACCAATTGA AGAAATATGA CTTTCAAAAA CCTTCTAAAA TTCCTTATCT TGAAACGACT | 5100 |
| GGTATGCCTA CAAGAATTCT CCTTAGAAAG CGTCGATTCA AGTGCTATCA CTGTTCAAAA | 5160 |
| ATGATGGTCG CTGAACTTC TGATGACGTA CAGTCATATT TCTTCTCTTT TTATTATATC | 5220 |
| ACAGTTTAA ATCTAGCTTT ACTAGATTCA CCGCTACTAT CTATTATTTC GGAAAAAGA | 5280 |
| CGAAAAACC TGAGAATCAT CTCAGGCTTG GTCATTAAAT TTTTCTCTCA ATATCGAAAA | 5340 |
| GTGGAGAAAG TGGTCTTTT TCATGAATAC GTACGATAGC ATCCCTTAGG AGATGAGCGA | 5400 |
| TTGAAATCTG CTCAATCTTA TCAATCAAAC GCTCTTCTGG CAGATAGATG GTATCCAAAA | 5460 |
| CAACCAATTT CTTAATAGCT GATTTTGGGA TATTGTCCGT AGCAGGACCA GAAAGAACTG | 5520 |
| GGTGCGTACA GCTTGCATAG ACTTCAACAG CACCAGCTTC CGCAAGAGCA TCTGCCGCAT | 5580 |
| GACAAATCGT TCCAGCGGTA TCAATCATAT CATCAATCAA GATACAAGTC TTGCCTTCAA | 5640 |
| CCTTACCGAT GATATTCATA ACTTCACTAG TATTCATCTT ATCAACGCTA CGACGTTTAT | 5700 |
| CAATAATAGC GATAGATGTT TTCAAAAATT CTGCCAATT ACGAGCACGA GTCACCCCTC | 5760 |
| CATGGTCCGG GCTGACAACC ACATAGTCAG AACCAACCAT ACCACGACGC TCAAAATAAT | 5820 |
| CTGCAATCAG AGGAGCACCC ATCAAATGAT CCACAGGAAT ATCAAAGAAT CCTTGAATTT | 5880 |
| GCGCAGCATG CAAGTCGATG GTCAATAAAC GATCCACTCC AGCTACTTCA AGCATATTTG | 5940 |
| CGACAAGTTT TGAAGTGATT GGCTCACGCG CTCTCGCCTT TCTATCCTGA CGTGCATACC | 6000 |
| CATAGTAAGG CATGACAACA TTGACAGATT CTGCACTCGC ACGCTTCAA GCATCTACCA | 6060 |
| TAATCAAAAT TTCAAGCAGA TTGTCAATTA CAGGCGAACT AGTTGATTGT AAGATAAAGA | 6120 |
| CGTGTPTCCC ACGGATTGAT TCTTCAATGT TGACCTGAAT CTCTCCATCT GAAAATTGGC | 6180 |
| GAACACTTGA TTTCCCCAAC TCTATCCCAA TCTCTGCGC CACACGTTCT GCCAATTCTT | 6240 |
| TATTAGAAGA AAGGGCAAAC AGCTTTAAAT CAGAAAAAGA CATGATTTCC TCCGGTATAT | 6300 |
| ATGTATAACT TGTGCTTTT ACAAGATTTT CCATCTACCA TTGTAGCGCT TTTTGCCTA | 6360 |
| TTTTTCAATC AAAAATAAAA GAAGGGCACC ATATTTGTAC CCTTGCATCA TTCTTTTGAA | 6420 |

606
AAATATTCTA GGTCAATCAAC TCATTGTGTT TCTCAACAAA GCAATAAGCA TGATAAAAAC 6480
CATAGAGAGC AATAGCCGTA ACCACTGGAA TCGCTAAAGG CAACTCTGTT TCCAACTCCA 6540
CAAAAGGAGA GTTAAACAAG AAGTGAGTTC CCAAGGCTAA ACCTAGAAAA ATAAGGCCCT 6600
GTTTCTTGCC AACCTTCTGT CCTTTATAGG CTCTGTAAAG CAAGTAAACA CCTACTACAG 6660
CTAGACCTGA AAAAGTCCAG TGAGAGGCAA TTCCTGAGAT GATACGCTCT AAAATTCGCG 6720
AAATAGTAAA GTCAAAGCCC TCTGGCAAAT CCGTACGAAT ATAACCAATA TCCTTAATCA 6780
TTTGGAATCC CAAACCGGAA GCAATTCCAA GTAAAAACAA AGATTTTAAT TTTCGCACAG 6840
GAATCAAAGC CAAAACAAAA ACAAGTGACA ATAATTTCAA GGGTTCTTCT ACCAAAGGAG 6900
CCGCAATAGC ACTTTCAAAG GCATTTAAAA ATGGACTATC TGGGAAAAGA ACCCCAGTA 6960
AATCATGGAT ATAAGTATTA GCAAACTAG ACAACCAGCC TGAAAGGAAC ATCCCTCCCA 7020
ATAAAGACAG AATCAAAACC TTCTTTGGCA ATTCCCATTT TTCCAATAC GGAAGAGAAA 7080
ATAAAGAGCC GGAATCATGT AAAAGAGAGC TAGAAAGATA GAAACTCCCA TTAGTCCATA 7140
TTCCGCACCT GACCTCGAAC CGTCCGTATA GTAGATGGTT TCATACTGTA AACCAATACA 7200
TAGCAATAAA ATAAAAATAA ATAAATATT GCTTTTCTTC ATACACTTTC TTTCTAAATG 7260
AAGTATTTAT AATTCTACGA CTGTCATACT TCCTGTATCA ACATTGTAAA TGGCACCAGA 7320
GATAATGACA TCGTCTGGTA TTAGGGGAGA CTCGATAAGC AGTTGCATAT CCTCGCGTAC 7380
ACTCTCTTCT ATATCTTGGA AGGGCAAGAA GTCTGTGCT GACACATCGA CACCCAATTC 7440
TTCCTTCAAA TACTCCTGAA AAGGTTCAAT TTCAAAGGTC TGAGCACCAC AGTCTGTATG 7500
ATGCAATACC ACAATTTCTC TTGTCCCAT TTGTTGCTGG GAAATAACTA GAGAACGAAT 7560
CATATCCTCA GTCACGAC CACCTGCATT CCGCAAAATA TGAGCATCCC CAAGTGCCAA 7620
ACCTAGAGCT TGCGCAACGT GCAAACGTGA GTCCATACAG GTCACAATGG CTACTCTGGT 7680
TTTAGGTTTA AGTGGCAGAT TTAAGTGGCC ATGTAGGGCA ACATAAGCCT GATTGGCTTG 7740
CATAAACTGT TCAAAATACG ACACGATTCC CTCCTTGAAA ATTTGATAGT CAAATATTTT 7800
TCCTATCTTA TCATTTTAA GAGAATTGTT CACGGATTAT GCAAAGACCT TTTTCAAGAC 7860
TTCCTGAATC GTTGTCACGC CAATGACCTG AATTTCCTTA GGCAGAGTGA TTCCTGTCAA 7920
GGAATTCTTA GGTACATAAA TCTTAGTAAA GCCCAGTTTA GCAGCTTCGT TGATGCGTTG 7980
CTCAATACGA TTCACGCGCC GAATCTCTCC TGTCAAGCCC AGTTCTCCGA CAAAACATTC 8040
CTGAGGATTA GTTGGCTTGT CTTTGTAGCT CGAAGCAATA GCAACTGCAA CAGCCAAGTC 8100
AATCGCAGGT TCATCCAATT TAACACCACC AGCAGATTTG AGATAGGCAT CCTGATTTTG 8160
CAAGAGAAGC CTGCCCCTT TTTCCAAAAC AGCCATAATC AAGCTAGCAC GGTTAAAATC 8220

607

| | |
|---|------|
| AAGTCCTGTC GTAGTACGCT TGGCATTTCC AAACATGGTC GGTGTTACCA AAGCCTGAAC | 8280 |
| CTCCGCCAAA ATCGGACGCG TCCCTTCCAT GGTTACAACG ATGGAGGAAC CAGTCGCCCC | 8340 |
| ATCCAAACGC TCTTCTAGGA AAACCTGACT CGGATTGAGT ACCTCAACCA AGCCGCCCGA | 8400 |
| CTGCATCTCA AAAATCCCAA TCTCATTAGT GGAACCAAAA CGATTTTGA CCGCTCTCAA | 8460 |
| AATACGAAAG GTGTGGTGAC GCTCCCCTTC AAAGTAAAGC ACCGTATCCA CCATATGCTC | 8520 |
| CAACATACGA GGCCAGCCA AGGTTCTTC TTTGGTCACA TGACCTACGA TAAAGATGGC | 8580 |
| AATGTTATTG GTCTTGCCA ACTGCATGAG TTCAGCGGTC ACTTCACGCA CCTGAGAAAC | 8640 |
| AGACCCCTGC ACCCTGAAA TCTCAGGAGA CATGATGGTC TGGATGGAAT CAATAATGAG | 8700 |
| AAAGTCTGGC TGGATACGCT CCACTTCTGC ACGAACACTC TGCATATTGG TCTCTGCATA | 8760 |
| GAGATAAAAC TCACTATCAA TATCACCTAA GCGCTCTGCA CGTAGTTTAA TCTGCTGGGC | 8820 |
| AGACTCTCC CCACTGACAT AGAGAACTGT CCCCCTTGG GACAACTGGG TTGAGACTTG | 8880 |
| TAGGAGAAGA GTTGATTTC CAATCCCAGG ATCCCCACCG ATAAGGACGA GACTTCCTGG | 8940 |
| TACCACTCCG CCTCCAAGCA CACGGTTGAA TTCCTCCATC TCCGTCTTG TTCGATTGAC | 9000 |
| ATTGATGGAA GTCACCTCAG CTAGTTTCAT GGGCTTGGTT TTCTCACCTG TCAAGGACAC | 9060 |
| ACGCGCATTC TTAACCTCG CAACCTCAAC CTCTTCCACA AAAGAAGACC AAGACCCACA | 9120 |
| GTGCGGCA CGTCCAGAT ATTTAGGGGA ATTATACCCA CAATTTTGAC ATACAAATGT | 9180 |
| CGCTTTTTC TTTGCGATGA CAAACCTCT TCTATATCTC TAACTCACAC TCAATCACTT | 9240 |
| GGCAAAAATC AATCTTCTCA TTTGGCACAA ACTGGCGCAT GAGCATTCGA TGAGCAACAA | 9300 |
| CTACCACAGT CTGATGTTCT CGATACTTAG ACATACATTC TAGAAACCGA GACTTCATTT | 9360 |
| CCGTAGCTGT CTCATATTGA ATAGGACTAT TAGGAAGCAA CTCCCCCTTG TTTTCTAAAA | 9420 |
| ACAGTCTTCT AGCTGTTTCA AAGTTTCTA TTCCTGTTT ATAGACCTGC CATTCATGTA | 9480 |
| ATAAAGGCTC TACTCTTAAA GGAAGACCCG TAGCACAGAC CACATACGAA GCCGTTTCTA | 9540 |
| AAGCTCTTGT GACTGCAGAA GATACGATTA TTTCAGCTGA CGAGAGTAAA GGATTTTGC | 9600 |
| TCAATTTCTG GACTTGCTGC CGTCCATCT CAGACAAGGG TGCCAAATCT ATCCCAAATC | 9660 |
| CTATATAAGA ACGCTCTCT AACTCACGGT AATCTGGCTC CCCATGACGT ACAAAGATAA | 9720 |
| TCTTCATTCT AGTGCCCTGT CGATCCAAAT CCACCAGTTC GAACGCCATC AGCTGCATCT | 9780 |
| CCATCTGCAA TTAAGAAAGT AGCAAAAACA GCCTGGACAA TACGCTCCCC AACTTCAAGA | 9840 |
| ACAACCTCTT GGTCTGTGAT ATTCTTCATC TGCGCAAAA TATGCCCTTC ATTTCCAGGA | 9900 |
| TTTCATAAT AATCCCCATC AATGACTCCA ACTGAGTTAA TTAAAACCAA GCCCTTCTTA | 9960 |

608

| | |
|---|-------|
| CGAGGATTTG AAGAACGATC ATAGAGGTAG AGAACCTCAG TCGGCTGCAT ATAAGCCTTA | 10020 |
| ACCCCTGTGCG GAACCAAGAC AATCTCTCCT GCGCAACAA CTGTACGCAC AGCAACCTTT | 10080 |
| AAGTCGTAAC CAGTCGCATG CGCTGTCTCA CGCTTGGGCA ATAAATTTTC ATCTGTAAAA | 10140 |
| CTCGAAACCA ATTCAAAACC ACGAATTTTC ATAATTTTCT CTTTCTATT ATCATTATT | 10200 |
| CTAGATTATT CTATACTTAT TTA | 10223 |

(2) INFORMATION FOR SEQ ID NO: 74:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 16535 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:

| | |
|---|------|
| TGGTCTGTC CTTATCGGCG CCTTGTCTTG CTGCCATGG CTACACCAAC TATCTCATCC | 60 |
| GACGAAAGTA CACCAACCAC TAACGAACCC AACAACAGAA ATACAACCAC CCTTGCCCAA | 120 |
| CCTCTTACTG ATACAGCAGC TGGCTCTGGT AAGAACGAAA GTGATATTTT TACACCTGGA | 180 |
| AATGCAAACG CTTCCCTAGA GAAACAGAA GAAAAACCTG CTGCAAGCCC AGCCGATCCA | 240 |
| GCACCACAAA CTGGACAAGA TCGTTCAAGT GAGCCAACTA CTTCTACTAG TCCAGTAACA | 300 |
| ACTGAAACTA AGGCAGAAGA GCCCATCGAA GATAACTACT TCCGTATCCA TGTCAAAAAA | 360 |
| CTTCTGAAG AAAACAAGGA TGCTCAAGGA CTATGGACTT GGGACGATGT TGA AAAACCA | 420 |
| TCTGAAACT GGCCAAACGG AGCTTTGTCC TTCAAGGATG CCAAGAAAGA TGA CTACGGC | 480 |
| TATTACCTAG ATGTCAAATT AAAGGGAGAA CAAGCCAAGA AAATTAGCTT CCTCATCAAC | 540 |
| AATACAGCTG GAAAAATCT AACCGGCGAT AAATCTGTAG AAAAAGTAGT TCCAAAAATG | 600 |
| AACGAAGCTT GGTTAGACCA AGATTACAAG GTTTTCTCTT ACGAGCCACA GCCTGCAGGA | 660 |
| ACTGTTTCGG TCAACTACTA CCGCACAGAT GGCAACTATG ACAAGAAATC TCTCTGGTAC | 720 |
| TGGGGAGATG TGAAAAATCC AAGTAGCGCT CAATGGCCTG ACGGAACAGA CTTTACGGCT | 780 |
| ACAGGCAAAT ATGGCCGCTA TATCGACATT CCTCTTAATG AAGCCGCAAG AGAATTTGGA | 840 |
| TTTTTATTAC TAGATGAGAG CAAACAAGGA GACGACGTGA AAATCCGTAA AGAAAATTAT | 900 |
| AAGTTCACAG ATTTGAAAAA TCATAGCCAA ATTTTCCTAA AAGACGATGA TGAATCGATT | 960 |
| TACACAAATC CATACTATGT CCATGATATC CGTATGACAG GAGCCCAACA CGTAGGCACT | 1020 |
| TCTAGCATTG AAAGTAGCTT TTCAACACTT GTCGGTGCTA AAAAGAAGA TATCCTCAAA | 1080 |
| CACTCCAACA TCACTAATCA CCTAGGAAAC AAGGTAAC TAACCGATGT TGCAATCGAT | 1140 |

609

| | |
|---|------|
| GAAGCTGGTA AGAAAGTGAC CTACAGCGGA GATTTCTCTG ACACAAAACA TCCTTATACT | 1200 |
| GTTAGCTACA ATTCGACCA ATTCCTACC AAAACAAGCT GGCGCCTGAA AGATGAGACA | 1260 |
| TACAGCTATG ATGGCAAACCT GGGAGCTGAC CTAAGAAG AAGGAAAACA AGTTGATTTG | 1320 |
| ACCCTTTGGT CACCAAGTGC TGATAAGGTT TCTGTTGTTG TCTACGACAA GAATGACCCT | 1380 |
| GACAAAGTAG TTGGAACGT CGCTCTTGAA AAAGGGGAAA GAGGAACTTG GAAACAACT | 1440 |
| CTAGACAGCA CAAACAACT CGGAATCACA GATTTCACTG GCTACTATTA TCAATACCAA | 1500 |
| ATCGAGCGTC AAGGTAAAC TGTTCTTGCA CTCGATCCTT ACGCTAAATC TCTTGCTGCT | 1560 |
| TGGAATAGCG ACGATTCCAA GATTGACGAT GCCCATAAAG TGGCTAAAGC CGCCTTGTA | 1620 |
| GATCCAGCTA AACTCGGACC TCAAGACTTG ACTTATGGTA AGATTACAA TTTCAAGACT | 1680 |
| CGTGAAGACG CCGTTATCTA CGAAGCTCAT GTGCGTGATT TCACTTCAGA TCCTGCCATT | 1740 |
| GCAAAGACT TGACCAAACC ATTTGGGACT TTTGAAGCCT TCATTGAAAA ACTAGACTAT | 1800 |
| CTCAAAGACT TGGGTGTAAC CCATATCCAG CTCCTTCCAG TCTTGTCTTA CTACTTTGTC | 1860 |
| AATGAATGA AAAACCATGA ACGCTTGCTT GACTACGCTT CAAGCAACAG CAACTACAAC | 1920 |
| TGGGGATATG ACCCTCAAAA CTACTTCTCC TTGACTGGTA TGTACTCAAG CGATCCTAAG | 1980 |
| AATCCAGAAA AACGAATCGC AGAATTTAAA AACCTCATCA ACGAAATCCA CAAACGTGGT | 2040 |
| ATGGGAGCTA TCCTAGATGT CGTTTATAAC CACACAGCCA AAGTCGATCT CTTTGAAGAT | 2100 |
| TTGGAACCAA ACTACTACCA CTTTATGGAT GCCGATGGCA CACCTCGAAC TAGCTTTGGT | 2160 |
| GGTGGACGCT TGGGGACAAC CCACCATATG ACCAAACGGC TCCTAATTGA CTCTATCAAA | 2220 |
| TACCTAGTTG ATACCTACAA AGTGGATGGC TTCCGTTTCG ATATGATGGG AGACCATGAC | 2280 |
| GCCGCTTCTA TCGAAGAAGC TTACAAGGCT GCACGCGCCC TCAATCCAAA CCTCATCATG | 2340 |
| CTTGGTGAAG GTTGGAGAAC CTATGCCGGT GATGAAAACA TGCCTACTAA AGCTGCTGAC | 2400 |
| CAAGATTGGA TGAAACATAC CGATACTGTC GCTGTCTTTT CAGATGACAT CCGTAACAAC | 2460 |
| CTCAAATCTG GTTATCCAAA CGAAGGTCAA CCTGCCTTTA TCACAGGTGG CAAGCGTGAT | 2520 |
| GTCAACACCA TCTTTAAAAA TCTCATTGCT CAACCAACTA ACTTTGAAGC TGACAGCCCT | 2580 |
| GGAGATGTCA TCCAATACAT CGCAGCCCAT GATAACTTGA CCCTCTTTGA CATCATTGCC | 2640 |
| CAGTCTATCA AAAAAGACCC AAGCAAGGCT GAGAACTATG CTGAAATCCA CCGTCGTTTA | 2700 |
| CGACTTGGA ATCTCATGGT CTTGACAGCT CAAGGAACTC CATTTATCCA CTCCGGTCAG | 2760 |
| GAATATGGAC GACTAAACA ATTCGCTGAC CCAGCCTACA AGACTCCAGT AGCAGAGGAT | 2820 |
| AAGGTTCCAA ACAAATCTCA CTTGTTGCGT GATAAGGACG GCAACCCATT TGACTATCCT | 2880 |

610

| | |
|---|------|
| TACTTCATCC ATGACTCTTA CGATTCTAGT GATGCAGTCA ACAAGTTTGA CTGGACTAAG | 2940 |
| GCTACAGATG GTAAAGCTTA TCCTGAAAAT GTCAAGAGCC GTGACTATAT GAAAGGTTTG | 3000 |
| ATTGCCCTTC GTCAATCTAC AGATGCCTTC CGACTTAAGA GTCTTCAAGA TATCAAAGAC | 3060 |
| CGTGTCACC TCATCACTGT CCCAGGCCAA AATGGTGTGG AAAAAGAGGA TGTAGTGATT | 3120 |
| GGCTACCAAA TCACTGCTCC AAACGGCGAT ATCTACGCAG TCTTTGTCAA TGCGGATGAA | 3180 |
| AAAGCTCGCG AATTTAATTT GGGAAGTCC TTTGCACATC TAAGAAATGC GGAAGTTTGT | 3240 |
| GCAGATGAAA ACCAAGCAGG ACCAGTCGGA ATTGCCAACC CGAAAGGACT TGAATGGACT | 3300 |
| GAAAAAGGCT TGAAATTGAA TGCCCTTACA GCTACTGTTC TTCGAGTCTC TCAAAATGGA | 3360 |
| ACTAGCCATG AGTCAACTGC AGAAGAGAAA CCAGACTCAA CCCCTTCCAA GCCTGAACAT | 3420 |
| CAAAATGAAG CTTCTACCC TGCACATCAA GACCCAGCTC CAGAAGCTAG ACCTGATTCT | 3480 |
| ACTAAACCAG ATGCCAAAGT AGCTGATGCG GAAATAAAC CTAGCCAAGC TACAGCTGAT | 3540 |
| TCACAAGCTG AACAACCAGC ACAAGAAGCA CAAGCATCAT CTGTAAAAGA AGCGGTTCCA | 3600 |
| AACGAATCGG TAGAAAATC TAGCAAGGAA AATATACCTG CAACCCAGC TAAACAAGCT | 3660 |
| GAATTCCAA ATACAGGAAT CAAAAACGAA AACAACTCC TATTTGCAGG AATCAGCCTC | 3720 |
| CTTGCGCTCC TTGGTCTCGG TTTCTTACTA AAAAATAAAA AAGAGAACTA AACTAGCCCT | 3780 |
| CCTATAGAAA AATCCCCCA GCATTATAGC TCGGGGGATT AATTTTGTGTA CAATATTTGT | 3840 |
| TGTCCTAATA AACTTGATTA GGATTTTTTA TTAAGCCTCT TTCATAGCAA AATAAGCTCG | 3900 |
| TACTTTGGGT GCAACTTGTG TTCCGAAGAG TTCAATAGCT CTCAGAACCT GGTCAAGAG | 3960 |
| CATAGAACCA AGCGGTAGAT GAAGCATGAA GCGGTCCAAT CCTAAATCCT CTATCATGCG | 4020 |
| AATCAATTTT TCGGCCACCT GATCTGGATT GCCAACAAAC ATGGCGCCAT TTGGCCCTAC | 4080 |
| CTGCTCCAAA TATTGCTCAT AACGCAATC CTGCCAGTGC GGACGGTCTT TGGAAATAGC | 4140 |
| ATCCACCACT TGCTTAGTCG GATGGAATA ATCTTTCACC GCCTGCTCAC CATCTCCGC | 4200 |
| AATCCACCCC CAAGAATGGG CTCCCACTT CAAGTCTTTG TCAGCATGGC CCCTTCGCTT | 4260 |
| CCAATCTCAC GATAAGCCTG AATCAACTTT TTAATAAAC GTGGATTACC ACCAATAATA | 4320 |
| GCATATACAA TCGGTAGACC AGCCTGAGCA ATCTTCACTG TTGATTGAC ATGACCACCT | 4380 |
| GTAGCTATCC ACAAGGGCAA TTTGTCTGA ACTGGACGAG GATAAACTTC TTTACCAGCA | 4440 |
| ATCGTTTGAG TCAATCGACC TGCCAGTCT AACTTGGTCT TTTCAATTGAC TAACTGAAGC | 4500 |
| AAGTCTAATT TCTCATCAA AAGAGAGTCG TAGTCTTTCA AGTCATAACC AAACAGAGGG | 4560 |
| AAAGATTCCG TGAAAGAGCC CCTTCCAGCC ATAATCTCCG ATCGTCCATT TGACAAAGCA | 4620 |
| TCGATAGTGG CATACTGTTG GAACAAACGA ATCGGGTCCA TGCTTGACAG AATGCTGACT | 4680 |

611

| | |
|--|------|
| GCACTGGTCA AACGGATTTT CTTGGTATTG ACTGCCCCAG CGGCCAGAAC AATCTCTGGG | 4740 |
| GCTGATACTG CAAAATCCGC CCGATGGTGC TCACCAATCC CATATACATC CAAACCAACC | 4800 |
| TTGTCAGCCA GCTCAATCTC TGCCACCAAC TGGCGAATGC GTTCAGCATG ACTGTAAGTT | 4860 |
| TGTCCAGTCC CTTCAAGCTC CGTTATTTCC CCAAATGTTG AAATTCCCAA TTCTACCATT | 4920 |
| GTGATTCTCC TTATCTATCT CTGTACTTCA ATTTGAAAAA TTATTCTAAC ACGAATCTTG | 4980 |
| AGTACAAGCA ACCGATTTGC TCATTAGAAA AAGCCTAGAT AACTAGACTT TTTTAGCTTA | 5040 |
| TTCTACCGTT ACTGACTTGG CAAGGTTACG TGGTTGTCC ACATCGAGGC CACGGTGGAG | 5100 |
| GGTTGCAAAG TAAGCGACTA ATTGCGTTGG TACGACCATT GAAATTGGTG AGAGGTATGG | 5160 |
| ATGTACGGTC GTAAGGACGA TATCGTCGGT ATCTTTGGCT ACATTCTCTT CTGCGATAGT | 5220 |
| GAGGACTTTG GCACCACGGG CTGCGACCTC TTGGATATTT CCACGAGTAT GATTGGCAAG | 5280 |
| AACTGGATCT GACAAGAGAG CAAAACAGG CGTTCCTTCT TCAATCAAGG CAATGGTTCC | 5340 |
| GTGCTTGAGT TCTCCTGCAG CAAAGCCTTC AACTGGATA TAAGAAATCT CTTTGAGTTT | 5400 |
| GAGACTTGCT TCCATGGCTA CGTAGTAATC TTGACCACGT CCGATGTAAA AGGCGTTACG | 5460 |
| AGTTGTTTCA AGAAGTTCAC GAACCTTGAC TTCAATGGTT TCTTCTCTG AAAGAGTTGA | 5520 |
| TTGATAGAC TGAGCTACGA TTGACAATTC ATGAACCAGG TCAAAGGCTT GCGCTTTAGC | 5580 |
| ATTACCATTT GCTTCTCCGA CTGCTTTTGC AAGGAAGGCA AGGGCTGCGA TTTGCGCTGT | 5640 |
| ATAGGCTTTA GTTGATGCCA CGGCAATTC AGGACCTGCG TGAAGGAGCA TGGTATAGTT | 5700 |
| GGCTTCACGT GAGAGGGTTG AACCTGGAAC GTTGTCACT GTTAAGCTTG GAATCCCAT | 5760 |
| TTCAATTAGCC TTGACCAAAA CTTGACGACT ATCCGCTGTT TCACCAGATT GGCTGATAAA | 5820 |
| GATGAAGAGT GGTTCCTTGC TGAGAAGTGG CATACCGTAG CCCACTCAG ATGAGATTCC | 5880 |
| AAGTTCAACT GGTGTATCTG TCAATTCTTC CAACATTTTC TTAGAAGCAA ATCCTGCATG | 5940 |
| GTAAGATGTT CCAGCTGCAA GGATGTAGAT GCGGTCTGCG TCTTGAACAG CCTTAATGAT | 6000 |
| ATCTGGGTCT ACGACAACTT GACCAGCCTC ATCTGTGTAG GCTTGGATGA GTTCCGCAT | 6060 |
| AACAGTTGGT TGCTCGTCAA TTTCCTTGAG CATGTAGTAA GGGTAAGTTC CCTTACCGAT | 6120 |
| ATCTGACAAG TCAAGTTCAG CAGTGTAGCT AGCACGCTCA CGACGATTTT CATCATAGTC | 6180 |
| TTGAACCTTC AACTATCAG CTTGACGAT TACCAACTCT TGGTCATGGA TTTCCATGTA | 6240 |
| TTGGTTAGTT TCACGAATCA TAGCCATGGC GTCTGAGCAG ACCATGTTAT AGCCTTCTCC | 6300 |
| AAGACCAATC AAAAGTGGTG ATTTATTTTT AGCTACGTAG ATGACTTCAG GATCTTGTGA | 6360 |
| GTCAACCAAG GCAAAGGCAT AAGAACCACG GATGATGTGA AGGGCTTTTT TGAAGGCTTC | 6420 |

612

| | |
|--|------|
| AAGAACTGAG AGCCCTTCTT CTTCCGCAAA TTTTCCAATC AAATGAACGG CTATTTTCAGT | 6480 |
| ATCTGTCTGC CCCTTGAAGT GGTGACCTGC AAGGTATTCT TCCTTGATTT CAAGATAGTT | 6540 |
| CTCAATCACC CCATTATGCA CCAAGACAAA ACGTTCCGTC TCAGAGCGGT GTGGGTGAGC | 6600 |
| ATTGTCCTCA GTTGGTTTTT CGTGAGTAGC CCAACGAGTA TGTCCGATAC CAGTTGTTCC | 6660 |
| CTCAACACCA GCTGTCTTGG CAGACAATTC TGCAATACGA CCAACCGCCT TCACCAAATG | 6720 |
| GTTATCAGCA CCATCTAGGA CAAAAATTCC CGCAGAATCA TAGCCACGGT ATTCAGGCTT | 6780 |
| TTCAAGCCCT TGAATCAAAA TATCAGTTGC ATTTGTGTTT CCAACAACAC CAACAATTCC | 6840 |
| ACACATAGTA TATACGACAC AGGCAAGCTG TGCTTTCTCC TTAAAATTGG TATAGTCTAA | 6900 |
| TTCATCTTTT ATAGAATCAG CAAAAACAGT ATATACTTGT TTCTTTCACT TGTCAAGAGT | 6960 |
| AAAAATTGGT ATAGTTCAAA TTAAGCTCCT GTAAGCATAA AACTCTGAC CGATTGGGAT | 7020 |
| AATCAGTCAG AGTCCTTTTT AAAATCCATT ATTATCGCTT AATTCITTGA ACCAGTGGCC | 7080 |
| TGATTTCTTC AGACGACGTT CTTGCGTTTC CAAGTCTAAT TCGACCAAAC CATAGCGATT | 7140 |
| TTTATAGCTG TTGAGCCATG ACCAGCAGTC AATAAAGGTC CAAATCAAGT AGCCCTTACA | 7200 |
| GTTGGCACCA TCTTCAATGG CACGGTGAAG TTCACGAAGA TGACCTTTTA CAAAGTCAAT | 7260 |
| ACGGTAATCA TCTTGAATCA TTCCATCTTG ACGGAATTTT TCTTCCCCTT CAACCCCAT | 7320 |
| ACCATTCTCA GTCAACATCC ACTCAATATT GCCATAATTT TCCTTGATAT TTTGGGCGAT | 7380 |
| GTCATAAATC CCTTGCTCAT AAATCTCCCA ACCACGGTGA GAATTGATTT TACGTCCAGG | 7440 |
| CATCACATAA GGCTCGTAAA AATGTTCTGG TAAGAGTGGG CTCTCTGGAT GCTTAGCAAA | 7500 |
| TCGAGGAGCC ATAACACGCA AAGGTTGATA GTAGTTCACA CCAAGGAAGT CCACCGTATT | 7560 |
| ATCACGAATG AGTTCCAAT CTTCCTCTGT AGCATCAGGT AAAAGACCGT GTTCATGCAA | 7620 |
| GATTCTTACC AACTCCTGTG GATAAGTCCC CAAGACAGAT GGATCTAAGA AAGATTGGGC | 7680 |
| CTGAAAAGG GCCGCAATAC GAGCTGCCTT GACATCAGCA GGATGCTGGC TACGTGGATA | 7740 |
| AGCCGGTGTC AAGTTGAGGA CAATCCCAAT CTTGGAATCA GGCAAAAGTT CATGGCAAGC | 7800 |
| CTTAACAGCC CGGCTGCTGG CCAATTGTGT ATGATAGGCT ACCTTAACAG CTGCCTCTGC | 7860 |
| ATCCACCTTA TGTGGATAAT GGGCATCATA AAAATAACCA AATCTACAG GAACGATGGG | 7920 |
| CTCGTTAAAG GTAATCCATT GATCCACTAA ATCTCCATAA GTCTCAAAAC AAAACGAGC | 7980 |
| ATAGTCTTCA TAGGCTGAGA CTGTCGCCTT ATTTTCCCAA CCATCACCAT CCTCTTGAAG | 8040 |
| GGCAAAAGGT AAATCAAAAT GATAGAGATT GACTAACAGA CGAATTCCTT TAGCCTTAAT | 8100 |
| AGCCTCAAAG ACCTTACGAT AAAAATCCAC ACCTTGAGTG TTGACTTTTC CACAGCCTTG | 8160 |
| TGGAATAATC CGTGACCACT GAATAGAAGT CCGAAAGGCT GTGTGACCAG TCTCTAACAA | 8220 |

613

| | |
|--|------|
| AAGCTCAATA TCCCGCTCCC AATTTTCATA AAAAGTCGAT GTCTTATCTG AACCAATCCC | 8280 |
| ATTATAGTAA CGATTTGGCT CCACTTGGAA CCAGTAATCC CAGAGATTGT CTCCCTTACC | 8340 |
| GTCACCAGCT ACACGTCCTT CTGTCTGCGG TCCAGAAGTA GAGGATCCCC AGACAAAATC | 8400 |
| CTTTGGAAAT CTAGCATAC ATTTACCTCT TTATCTACTC ATTTCTCCCA TTATACAGAA | 8460 |
| AAAACAAGGT AAAAAGTAGT TACATTTTTT CCTTGTTTTT CTTCTGATTA TAGTTTTTAT | 8520 |
| TTCTTGCTTA GGATTTCAAG CGTTTCAAGC ACGTTATCTG CATGAACCTC AATGGTGTCA | 8580 |
| CCAGTTGCCT TGATCTTAAC TTCTACAATG CCATCGGCCG CTTTTTTTACC AACAGTGATA | 8640 |
| CGGATTGGAA GACCAATCAA GTCATATCG CTAAATTTAA CACCGACACG TTCGTTACGG | 8700 |
| TCATCTGTCA AGACTTCATA ACCAGCTCCC ATCAAGCTTG CTTCAAGTTT TTCTGTCAAG | 8760 |
| GCTTGCCTT CTTCATCCTT GACATTGACA GTAATCAAAT GCACATCAA TGGTGCCAAT | 8820 |
| TCTTTAGGGA AATTGATTCC CCAAGCGTAA CGGTATTAC CTTTGGCGT TTTGTAAACA | 8880 |
| AAGAGGCGAG CGTGTGCTC CATCACTGCT GAAAGAAGAC GGCTGACACC GATACCGTAA | 8940 |
| CATCCCATGA TGATTGGCAC AGCACGACCA TTTTCATCCA AGACATCTGC TCCCATGCTT | 9000 |
| GCTGAATAGC GAGTCCGAG TTTGAAAATA TGACCGATCT CAATACCACG CGCAAAGTTA | 9060 |
| AGGACACCTT GTCCATCTGG GGAATTTCA CCCTCACGAA CTTACCGAT ATCCACATAT | 9120 |
| TCTGCAGTAA AATCACGGCC TGGGTTCA CAAGTCAAGT GGTAGTCATC TTCGTTAGCA | 9180 |
| CCGACAACG CATTGCGAAC ATCTGTACC TTACGATCTG CAATAATTTT AATATTCTCT | 9240 |
| GGCAAACCAA CTGGTCCAAG TGAACCAAAT CCTGCTTGAA CAACATTCGC CACTTCTTCT | 9300 |
| TCGCTAGCAA CGTCAAAGAA ATCTGCTCCC AAGTGATTTT TCAACTTGAC TTCGTTGAGT | 9360 |
| TGGTCATTTT CAACTAGAAG GGCTGCAACA AGCTCACCAT CTGCAATGTA GAAGAGGGTT | 9420 |
| TTAATCGTTT GTTCTTCTGG AACATTGAGG AAGGCTGCAA CTTTCATCAAT TGATTTAACA | 9480 |
| TCTGGCGTTG CAACACGAGT AACTTCTTCT TCAGCGACAA CACGGTTGCT TGGTTTGTAC | 9540 |
| TCGTTTGTG CCATTCTTAA GTTAGCTGCA TAGCTAGACT CACTTGAGTA AGCAATGGTA | 9600 |
| CTTTCACCAG AGACTATCCA TTTGAGCAAT TCTGCCTTGA TTTCTTCTTG CACTTCTGCA | 9660 |
| GGAATTTCTG CAAATGAGGC AACTGACTTG TCCAAGACAA CCCAGCGGTC AAGGTCTGTA | 9720 |
| CGAGCAGATG TAATGGCCAT AAATTCTTGG CTATCCTTAC CACCCATGGC TCCACCGTCA | 9780 |
| CCAATAATAG CCTGAAGTC TAAACCACTA CGAGTGAAAA TACGCTCATA GGCTGCTTTG | 9840 |
| TACTCATCAT AAACACTATC CAACTATCA TAGTTAGCGT GGAAACTATA AGCATCCTTC | 9900 |
| ATGATAAACT CACGTGTACG AAGAAGTCCA TTACGCGGGC GTTTTTCATC ACGATACTTG | 9960 |

614

| | |
|--|-------|
| GGCTGAATTT GATAAAGGTT GAGTGGCAAT TGCTTGTAAG ATTTAACAGA ATCACGGACA | 10020 |
| ATAGCTGTAA AGGTTTCTTC GTGAGTTGGA CCTAAGATAA AGTCTGATTT TTCACGGTTT | 10080 |
| TTTAGTTTGT AAAGGTCTTC ACCATAGGTT TCGTAACGAC CTGATTACAG CCACAATTCT | 10140 |
| GCACTAAGAA GGGCTGGAGC CAACATCTCA ACAGCACCAA TCTTTTCGAA TTCTTGCGC | 10200 |
| ATGATGTTTT TAGCTTTTTC AATCACACGG TTGGCAAGTG GTAGATAAGA ATAAACACCT | 10260 |
| GCTGAAACTT GGCGAACATA ACCAGCACGC AACATAAGAG CATGGCTGAT AACTTGAGCA | 10320 |
| TCGCTTGGA TTTTCGGAAG CGTTGGGATA GGCATTTTAC TTTGTTTCAT AATATTCCTC | 10380 |
| GATTATCTAA AAAAGAGTCG CATAATGTCA TTCCAAGTCA CAGCAATCAT CAAGACAACC | 10440 |
| ATGATGACCA CTCCGCCAA GGTGACATAG GTTTCAATTT CTTGTTTCAA TGGTTTGCGG | 10500 |
| CGGATGGCTT CTAGGATATT GAGCACAATC TTACCACCAT CCAAGGCTGG AATCGGAATA | 10560 |
| AGATTAAAAA TCCCAATATT GATGAAATC ATTGCCAAGA AGTACAAGAT ATTTTCAATT | 10620 |
| CCATTTTTAG CAGCATCACT ACTTGCCTTA AAGATAGCAA CAGGTCCACC CAACTTGTTT | 10680 |
| AAATCTGGTT GGAAATCAG ATTTTTCAGA GCTGAGAGAA TTCGGAGAGC TGAGTCAGCA | 10740 |
| GCAGTTGTAA AACCACCTAC AAACATGGAT AGAAAACTG ACTTAACCCC CGGTTGAACA | 10800 |
| CCTAGAAGGT AACGACCTTG ACTATCTTTG GGTGTAACAG TGACTTGTTT GTCCTCCCC | 10860 |
| TTTTCAGAAA TAGTCACATC CAAAGTCGGT GCCGCTTAT CTTTGGTTTC TGTTTCCACA | 10920 |
| GCTTGGATCA AGCTTTCCCA GTTGCTAACC TCATGTGAGC CAATCTTGGT AATTGTGCCC | 10980 |
| ATTTCTGGTA CTCCTACCTT GGCCAAGGCA CCTTGGGGCA TGATATGGAA CTGATTGGTA | 11040 |
| TCAACATCTC TGACACCACC CTGCATAAAG ATTAATAACCC AAAAAACAAC GACACCTAAG | 11100 |
| ATAAAATTGT TCATAGGACC TGCAAAATG GTAATCAGTT TGCCCCAGAT AGTCGCATTT | 11160 |
| TGATATTGAA CATCTAAAGG TGCAATCCGA ACCTCAGTAC CATCTGCTTC CACAACCGTT | 11220 |
| GCATCGTGAT CCACTGCAAA TGTTTTTTCT TCTTCCAGAA CCAATCCTTT GATAAAGAGC | 11280 |
| TTGTCTTCAA AATCAAACCTG GGTACCTGC ATAGGGAGGG CTGTTTGATC CAATTTTTTA | 11340 |
| CCTGAGAGAT TGATGCGTTT AACCTTACCA TCATCAGCAA GTGTCAAACCT AACAGGCGTT | 11400 |
| CCTGTCTTGA TTTCAGTTGT ATCATCACCC CAACCGGCCA TGCGGACATA GCCACCCAGA | 11460 |
| GGCAAGATTC GAATGGTATA GGCCGTTCCA TCCTTGCCAA TGTGAGCAAA AATTTTAGGT | 11520 |
| CCCATACCGA TGGCAAATTC ACGTACTAAA ATCCCTGATT TCTTGGCAAA GTAGAAGTGA | 11580 |
| CCGAACCTCGT GCACCACTAC AATAATCCCG AAAACCAGAA TAAAGGTTAA AATCCGAGC | 11640 |
| ATAGCGTTTC CTCCGTCTTT TGATTTAAAG AGTCCAAATA AGTGCATGAT TGGAAATACA | 11700 |
| AGCAACATAC TATCGAAACG ATCCAAAACA CCACCATGTC CAGGGATAAA TTTCCAGAA | 11760 |

615

| | |
|---|-------|
| TCCTTAACAC CAAAATGACG TTTGATCGAA CTTTCTAGTA AATCACCAAA TTGTCCAGCA | 11820 |
| ATGCTAAAGA AAATAGCAAA GACTGACATC TTGTAAATTC CATATGGAAG AGCAACTGTA | 11880 |
| CTGTCAACTA TCATAAGGAT AATGGTACT AAAATTGCTC CTAAAATACC ACCCAAGGCA | 11940 |
| CCCTCAAGGG TTTTATTAGG CGATACCCTT GGTGCTAACT TTCGTTTCCC ATAGTTCATC | 12000 |
| CCAACAAGAT AGGCACCACT GTCTGTCGCC CAGACGATAC ACAAGGCTAA GAGAGCCTTG | 12060 |
| TCCAAACCTG CAACACGAGC ATCTAGTAAA GCATTAAATC CAAAGCCAC GTAGAAGCTC | 12120 |
| ATAGCAAGAG GGAAAACCGC ATCCTCAATC GTATAAGACT TGCTAAAAAC GGTGTTTCCT | 12180 |
| AACATGATTG AAATCAAAAC ACTATAGGCA ACCACATTCC CATCAACTGG CAAAAAGTC | 12240 |
| AGGTAATCTT CCAAGGAAT GGTCAATGCA AAGGTTGCAA AGAGGGTCAA GAGGCCCTCC | 12300 |
| ATCGTCATGG TCTCTAGACC TCTCATCTTC AAAAGTTCAT GCATGGCTAG CATGGCTATG | 12360 |
| ATTCCGATTG CTATCTGAAG CAAGAGGCC CCAATCATTA AAATTGGTAG GAAAAAGCC | 12420 |
| AGGGCAATCC CTGCAACAA GGTTCCTTTC TGTAATCCT GGGTCATATT TCCTCCTAAA | 12480 |
| CTCCTCCAAA TCGGCGATGA CGACGATTAT AGGCAAGAAT AGCTTCCTGC AAGGCCGCTT | 12540 |
| CGTCAAAATC AGGCCATAAG GTGTCCGTAA AATAAGCTC ACTATAGGCT CCCTGCCATG | 12600 |
| GAAGGAAATT GCTCAAACGT AATTCTCCAC TAGTACGGAT AATCAAGTCT GGGTCTCGTA | 12660 |
| AGTCCTTAGG CAAATGCTGA GTAAAGAGAT AGTTACCAAT CAATTCCTCT GTGATGTCAC | 12720 |
| CTGGGTTGAT TTTGGCATCT AAAACATCCT GGGAAATCAA CTTAAGCGCC TGTGTAATCT | 12780 |
| CAGCACGTCC ACCATAGTTA AGAGCAAAAT TAAGAATCAA TCCTGTGTTG TTCTTAGTCA | 12840 |
| ATTCCTCAGC CTTGGTTAAA GCTTCAAAGG TTTGCTTAGG CAGGCGGTCT GTCTCCCCAA | 12900 |
| TCATTTGAAT CTTAACATTA TTCGCATGTA GTTCCGGGAC ATAATTATCA TAAACTCTA | 12960 |
| CTGGCAAATT CATGATAAAC TTGACTTCCT GATCTGGACG GGTCCAGTTT TCCGTAGAAA | 13020 |
| AAGCATAGAC CGTAATAACC TTGACGCCCA GTTTGTTGGC TGCCTTGGTC ACGGTTTGCA | 13080 |
| ATGCTTCCAT GCCCGCCTTA TGTCCAAAA CTCGCGGTTG CATACGTTT TTAGCCCAAC | 13140 |
| GGCCATTGCC ATCCATGATG ATGCCGATAT GAGCAGGAAC CTGTGTCGGA ACCTCTACTT | 13200 |
| CCACAGCCTT ATCTTTCTTA AAAAATCCAA ACATGATCTT ATTCCTATTC AAAATCTAT | 13260 |
| CGTTTCATTA TACCATATTT CCCCATTTTC TTCTATCACT AAGCTATTTA TTCTCAGGCA | 13320 |
| CCAAGCCCAT TTTTCAAAAA AATAAGCCGC CTGATTGGGC GACTTTATTT TTATAGGGAG | 13380 |
| ATTATTATGA AAAAGTTTAA GGAGTTTAA TTAAGGTCTT CTTAACTTAT GAACTTAGTG | 13440 |
| TACACTCCCT AGCTTAAAGT TTCCTTAAAGT ATTTTAAAA ATCAAATTTT TCCATTTCTC | 13500 |

616

| | |
|---|-------|
| CTGCCAATTT TTCTTGATA AACGTGTTG ATAGAGTTCC ATTCGGTCTT CATTTTCTAA | 13560 |
| GAAATGAGGA GTTGGACGAA CTTGAAAATT CAAAATATCC TCCAAACCAT AAGGTACATA | 13620 |
| GAGTTCAAAA TCTAATCTT CATTCAGCG CAGTCCAAC TCCGTACACC GTTCTGGATA | 13680 |
| CTTACTCATA GCATCACGAG AACTGGTATA GGAAGCAGTG TGAGGACTGT GCTGATGCAT | 13740 |
| ATAGACCTGA TTTTCAATT CCCACTGGTA CTGAGGAAAA TCCTCTCTCA GCTTTTCTC | 13800 |
| CAGTAATAAG GTTCTCTCAT AAGAAAAATC TGGATCAAAG AAAATCACAT CTATATCTGT | 13860 |
| TTCATGATCA AAAGGGGATT TGTCTGACAA AAGATTCAG ATGAAATTC TGACAGAACC | 13920 |
| TGCTGCCAAC CACGAGTCTT TCAAACCAAG GTCTCGGATG ATCGTCAGAA TGGCCATCAT | 13980 |
| ATCTGGACTT TCTCTAAAAG CCTCTAAGAT TTCTTGCTTA TTTTCTACTG TATTCTAAC | 14040 |
| CTAAGTGCTC ATATGCCTTA GCAGTCGCCA CCCGTCCAGA CCGTGTCCGC ATGATAAAAC | 14100 |
| CTTTTGAAT CAAGTAAGGC TCATACATGT CTTCAACTGT CTCACGCTCT TCGGCGATAT | 14160 |
| TCACAGAAAG AGTTCCTAGA CCAACAGGTC CTCCACTGTA CATCTCAATC ATGGTGCGAA | 14220 |
| GGATTTTGTG ATCCACATAG TCCAAACCTT CATGGTCAAC ATCCAGCATA GTCAAAGCCT | 14280 |
| TATCGGTAAT AACATCATCG ATAACCCCAT TCCCCATTAT CTGGGCAAAA TCGCGCACGC | 14340 |
| GCTTGAGGAG ACGATGGCA ATACGAGGGG TTCCACGACT ACGTAGGGCC AACTCAGATG | 14400 |
| CTGCCTCATG GGTGATTTCC ATCTCAAAA TATCTGCCGT CCGCTCGACA ATTTCTGTCA | 14460 |
| AGTCAGCATG AGCATAATAC TCCATATGAC CTGTAATCCC AAAACGTGCC CGTAGTGGAT | 14520 |
| TTGAGAGCAT ACCAGCCCGA GTCGTCGCAC CAATCAAGGT AAAAGGAGGC AACTCCAAAT | 14580 |
| GAACACTGCG ACTGCCTTCA CCAGCCCAA TCATAATATC GATGTAGAAG TCCTCCATGG | 14640 |
| CACTATAAAG CACTTCTTCC ACTGACATGG GTAAGCGATG AATCTCGTCA ATAAAGAGGA | 14700 |
| CATCTCCAGG CTCTAAATCA TTCAAAATCG CTACCAAATC ACCCGCTTTT TCGATAACAG | 14760 |
| GACCAGACGT TTGCTTGAGA TTGACTCCCA GTTCATTGGC AATGACAAAA GCCATGGTTG | 14820 |
| TTTTCCCAAG CCCTGGAGGG CCAAATAAGA GCACATGATC CAGCGTTTCA TCCCGCATTT | 14880 |
| TAGCGGCTTC GATAAAGATC TGAAGTTGAT CCTTAACCTT ATCCTGACCA ATATATTCAC | 14940 |
| GTAAATACTG AGGACGGAGC GTGCGTTCTA CTAACCTCTC ATCACCATC ATCTCATTAT | 15000 |
| CTAAAATTCT ACTCATGGCT CTATTATATC AAAAAAACA AGCCACAAAC AAAAAAGCCA | 15060 |
| CCTGATTGGG TGACTCCTAA GTTTAGCACT TATGTGGTAT AATATTATAC GGCACCTCTA | 15120 |
| CACCGCCTAC GAAAGGAGGT GAGATAGCCC ATGATGGAAT TAGTACTCAA AACTATTATC | 15180 |
| GGACCAATTG TGGTCGGTGT CGTTCTTCGT ATAGTCGATA AATGGCTAAA CAAGGACAAA | 15240 |
| TAGTGTCAAA AAAGACCTCA AGCTTATTTG GTCGTGAGCT TGGGGTCTTT TCTAGCCTAT | 15300 |

617

GATATAGAAC TAGTACTCAA TTCCTTTTTA TTATCCCATTA GTTCACGAAT TTTGTCAAAA 15360
 CTTTACATTT TCTTCAACCG CTGTACGACA AGACGGTTAA GATTAAGAGA ACGTTAGGGA 15420
 TTCTATCAAT TTCATAGAAA TTTTGATTTC GTAAACGAAG AGACAATCTT ACATGTCCT 15480
 TCTCATTTAA TACGCCACTA CTAGACAAGC AAAATCATT TACAGTAGT TCCAGTCCTT 15540
 CAATTAACAG TCACTTACAA TCAAATTGAG TTTGAACTAG CTGAAGCGAC CACAGACCTA 15600
 TTTCTTAGTC ATATTCGCTA AAAAAATCCC CGCCAAAATC TCAAAAAGTC CCCGCCAATT 15660
 CCCCAGCAA AATCCGAAAA ATACCGAAAA ATATCGAAAA ATTATTTTGA GAATAGTCCC 15720
 AAAAACTCTG AAATAGAGCT AAAAACTCC ACCTGATTCTG GTGGAGTTAA GGGAGATTAT 15780
 TATGAAAAAG AAAAGTTTAG GATTTTATTA AATAAAGTTA GGAGGTCTTT ATTTAATAAC 15840
 TACATGATAC AAGACGAAAC TTAAGACTAG CTAACTTTT CTAAAAATTTT ACTATTTTGC 15900
 AAAAAATTC TATCACCAGC ACCTCACCAA TCGAGTAGGG GATAATCTCT AGCCCCCTCTC 15960
 ACACCACCGT ACGTGCCGTT TGGCATACGG CGGTTCAACT AACTTTTAAC GCATGTCGTT 16020
 CAAGGTAATA ATCCAAACAC GAAACCAGTC CACGTTTTC CAGGACTGGT TTTGATATAG 16080
 CACGTTTAAG TACCGACTTC TGAGCTACTA ATTGATAATG GTCGCCCCAG CCAGATACCT 16140
 TATCTGCTAT CCATTTAGGA ACTCCTAACT TAAGCAATCC CCATAATCGT CTCGATTCTT 16200
 TCTTCCATTG CTTCCAGATA ATCACTCGTA GCGAGTACG CAAGCGCTCA TCTATGCTGG 16260
 CGACTATACT TTTCATATTT CCCAATGAGC AATAGTTTAT CCATCCTCGA ATAGACAAAT 16320
 TCAGTTGCTC AATACGTCTT GTTAGGTCTA TACTCCATTT CCTCTGTGTT AGTTTCTTCA 16380
 ATTTAACTT AAATCTCCGA AACTATCTT GATGTGGACG GCTTTTCCAA CCATCTGATA 16440
 ATTTCCAGAA CCCAAAACCT AGATATTTCA ACTCTCTTGG TCATGTTTAC TTTCAAACCT 16500
 AGCCGTTTCT CAATAAACGA CTGACTGAAT ACATC 16535

(2) INFORMATION FOR SEQ ID NO: 75:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 8136 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

CCAGAGCGTT GCGTCCGAAA GTCTATCCAG ACACGGCTCT TTA AAAACAA AAGGAGAAAT 60
 GATGCATACT TATTTGCAAA AGAAAATTGA AAATATCAA ACAACCCTAG GTGAAATGTC 120

618

| | |
|--|------|
| AGGTGGTTAC CGTCGTATGG TTGCGGCTAT GGCTGATTTA GGATTTTCAG GAACTATGAA | 180 |
| GGCTATCTGG GATGACCTCT TTGCCCATCG TAGTTTGGCC CAGTGGATTT ATTTGCTGGT | 240 |
| TTTAGGAAAGT TTTCTCTCTT GGCTGGAGTT GGTTTACGAA CATCGTATTG TTGACTGGAT | 300 |
| TGGGATGATT TGTAAGCTGA CAGGGATTAT CTGTGTAATC TTTGTATCGG AAGGTCGAGC | 360 |
| AAGTAATTAT CTTTTTGGCT TGATTAACTC TGTATTATAC CTTATTTTGG CCCTACAGAA | 420 |
| AGGCTTTTAT GGTGAGGTGC TGACGACACT TTACTTCACA GTCATGCAGC CAATTGGACT | 480 |
| TCTAGTTTGG ATTTATCAGG CACAGTTTAA GAAGGAAAAG CAGGAGTTTG TCGCGCGTAA | 540 |
| ACTGGACGGC AAGGGCTGGA CAAAGTATCT TTCCATTAGT GTGCTTTGGT GGTGGCCTT | 600 |
| TGGCTTCATT TATCAGTCTA TTGGTGCCAA TCGTCCCTAT CGTGATTCAA TCACAGATGC | 660 |
| AACCAATGGG GTAGGGCAA TCCTCATGAC AGCTGTTTAC CGTGAACAGT GGATATTCTG | 720 |
| GGCGGCTACC AATGTCTTTT CAATCTATCT CTGGTGGGA GAAAGCCTGC AAATTCAAGG | 780 |
| GAAATATCTA ATTTATCTCA TTAACAGTCT AGTTGGTTGG TATCAATGGA GCAAGGCAGC | 840 |
| TAAGCAGAAT ACTGATTAC TTAAGTAGGA AAAGATGTTT GAAAGTGCTG TTTTGAGATT | 900 |
| TCGATTAAAA CAGATATAGT TGATAATCAA GGATTTATAG TATGAAAAAG AGGATCGGCG | 960 |
| GGTCTCTTTT TGTGTGTA AAGATAAAAA ACTCAGTAAC CTAGAAATAA GACAACTGAA | 1020 |
| GCTTTACTCT ATATTCAATT TTTAGGAATG AGAAGGTCTA GATAAAATTG GACAACTTCC | 1080 |
| TGGTCTGTGA AATCTTGACC TTTTGTGAGC CACCAGGTCA ATGTCTCGAT AAAGTTGGAC | 1140 |
| ATGACCAAGT GTTGGAGGTA AGAAGTAGGC AGATTAGGGT GGGCTTCTTT TAAATTATCA | 1200 |
| GCTAGCACGG AATAGACATG GTGTTCTAGC TCTTTATGGA GTTGACGGAG GAAGTAGTCA | 1260 |
| TTTTTGAAA ATAGCAGACT GGTGATATGG TCTTGGTTTT TATGAAAATG GAGAAAGAGG | 1320 |
| TGGGCGAGGT AGTCCTCGGT TGAAATGGCT TGCTCTCTTT CAAAAGATG ATGGAAGAGG | 1380 |
| TAGCGGCAGA GCTGGTCCAG AAGAAGCTCC TTACTCTCAT AGTGACAGTA AAAGGTGGAT | 1440 |
| CGTCCCACAT CTGCGAGATC AATGATATCC TGAACAGTAG TGGCCTCGTA GCCCTTAGCA | 1500 |
| TTCAAAAGTT GTATAAAGC TTGATAGATG GCTTTTTTGG TTTTGCTGAT ACGGCGGTCA | 1560 |
| ATGTTAGTCA TATGGACACT TAAGGCAAAT TGTTCAAGAC TGAATAAAGC TGACGTTTGG | 1620 |
| CTTCTATCCT TTCTTTGAGT TTTAGTGGAT AATGATAATG AACAAAGTGT TCATAAATCT | 1680 |
| ATTATAACAA AGGAATGAGA AATATGAAGG CAAAATATGC TGTTTGGGTG GCTTTTTTCT | 1740 |
| TAAATTTGAC TTATGCCATT GTTGAGTTTA TTGCAGGTGG AGTATTTGGT TCTAGCGCTG | 1800 |
| TTCTTGCTGA CTCTGTGCAT GACTTGGGAG ATGCGATTGC AATTGGAATA TCAGCTTTTC | 1860 |
| TAGAAACAAT CTCCAATCGT GAAGAAGACA ATCAGTACAC CTTGGGCTAT AAGCGGTTTA | 1920 |

619

| | |
|---|------|
| GCCTGCTAGG AGCCTTGGA ACAGCTGTGA TTCTCGTAAC GGGCTCTGTT CTAGTCATTT | 1980 |
| TGGAAAATGT CACGAAGATT TTGCATCCGC AACCAGTCAA TGATGAGGGG ATTCTCTGGT | 2040 |
| TAGGAATTAT TGCGATTACT ATCAATCTGT TAGCGAGTCT GGTGGTTGGT AAGGGAAAGA | 2100 |
| CAAAGAATGA GTCTATTCTG AGTCTGCATT TTCTGGAAGA TACGCTAGGG TGGGTAGCTG | 2160 |
| TTATCCTGAT GGCGATTGTT CTTGATTTA CGGACTGGTA TATCCTAGAT CCTCTTTTGT | 2220 |
| CCCTTGTCAT TTCTTTCTTT ATTCTTTCAA AAGCCCTTCC ACGTTTTTGG TCTACACTCA | 2280 |
| AGATTTTCTT GGATGCTGTG CCAGAAGGTC TTGATATCAA GCAAGTAAAG AGTGGCCTGG | 2340 |
| AGCGATTGGA CAATGTGGCC AGCCTTAATC AGCTTAATCT CTGGACTATG GATGCTTTGG | 2400 |
| AAAAAATGC CATTTGCCAT GTTTGTCTAA AAGAAATGGA ACATATGGAA ACTTGTAAG | 2460 |
| AGTCTATTCG AATTTTCTTA AAAGATTGTG GTTTTCAAAA TATTACCATT GAAATTGATG | 2520 |
| CTGACCTAGA AACTCACCAA ACCCATAAGC GAAAGGTGTG TGACTTGGAA CGGAGTTATG | 2580 |
| AGCATCAACA TTAGAAAAAA GTGAAAAATA CTTGGGTACT ATCTTATTTG GAATAGAGTA | 2640 |
| ATTTCTTTAT TATTTAAATA TTTCAAAAAT TGGTAAGAGA AGAGCATTGT ATAACTCCA | 2700 |
| GATATATGAT TGTTAATGAT AAAAATTTTT CGATTAGATA CAAAATGCTT GACTTGGAGT | 2760 |
| CAACTCAAAG TTATATAATA AGATAAGTGA GTTAGAATAG CGTGAATTCA GTGAATGAAA | 2820 |
| TGAGAGGAGG TTAGCGTGTG AATATTAAAT CTGCCAGTGA TTTGTTGGGA ATTTCAGCGG | 2880 |
| ATACGATTTC GTATTATGAA CGGGTTGGTC TTGTGCCACC GATTACTCGT ACTGCTACTG | 2940 |
| GGATTTCGTG TTTTCAAGAT CAGGATATCG AAGCGCTGGA ATTTATTAAG TGTTTTCGTT | 3000 |
| CGGCGGGTGT CTCTGTAGAT AGTTTAGTTG ACTATATGTC GCTCTACCAA AAGGGAGATG | 3060 |
| AAACGAGAGA GGAGAGGCTT GGTATTTTAG AAGAGGAAAA GCAAAAATTA GAGGAGCGCT | 3120 |
| TGTCTCAGCT ACAGACAGCT TTAAATCGTT TAAATCTCAA AATTAACTT TATAAGGAAG | 3180 |
| GAAAATTTTA AATGAAATCA GCAGTATATA CAAAGGCAGG TCAGGTTGGA CTTGCTAGCA | 3240 |
| TTGAACGTCC GCAAATAATA GAAGCGGATG ATGTGATTAT TCGTGTGGTT CGTGCCTGCG | 3300 |
| TTTGTGGTTC AGATTTATGG AGGTACCGTA ATCCAGAAAC GAAAGCTGGA CACAAAAATA | 3360 |
| GTGGACACGA AGCGATTGGG ATTGTTGAAG AAGCTGGGGA AGCCATTACG ACGGTGAAAG | 3420 |
| CAGGTGATTT TGTGATTGTC CCTTTTACAC ATGGATGTGG TGAGTGTGAT GCCTGTCTTG | 3480 |
| CTGGATTTGA CGGTTCCTGC GACAATCATA TTGGCAATAA TTTGGGGGGT GATTTTCAGG | 3540 |
| CAGAATATAT TCGCTTCCAC TATGCAAACT GGGCGCTGGT TAAAATCCCT GGTCAACCTT | 3600 |
| CTGACTATAC AGAAGGGATG CTCAAGTCCC TTTTGACTCT TGCAGATGTC ATGCCGACAG | 3660 |

620

| | |
|--|------|
| GCTATCATGC GGC GCGTGTT GCAAATGTTT AAAAAGGGGA CAAGGTTGTT GTTATCGGTG | 3720 |
| ATGGGGCTGT TGGTCAATGT GCTGTCAATG CGGCTAAGAT GCGTGGAGCA TCACAAATTA | 3780 |
| TCCTTATGAG CCGTCATGAA GACCGTCAA AGATGGCTAT GGAGTCAGGT GCGACAgcTG | 3840 |
| TTGTTGCAGA ACGTGGTCAA GAAGGAATTA CCAAGGTGCG TGAAATCCTC GGTGGAGGAG | 3900 |
| CAGATGCAGC ACTTGAATGT GTTGGTACGG AGGCTGCTAT AGAACAGGCG CTAGGTGTTT | 3960 |
| TTCATAATGG AGGGCGTATG GGCTTTGTAG GAGTCCCACA CTATAATAAT CGTGCTCTTG | 4020 |
| GTTCGACATT TATGCAAAAT ATCTCTGTAG CAGGTGGGGC AGCTTCTGCT ACAACATACG | 4080 |
| ATAAGCAATT TTTACTAAAA GCCGTCCTTG ATGGTGATAT CAATCCAGGT CCGCTCTTTA | 4140 |
| CTTCAAGTTA TAAACTGGAA GATATCGACC AAGCCTATAA AGATATGGAT GAACGTAAGA | 4200 |
| CAATTAAGTC TATGATTGTA ATCGAATAAA AACGAATAG GAGTTTGTAGA ACTCTATTCTG | 4260 |
| TTTTTTTATGT TATCCTATTC TTGATTTAGG GTACTTTCTC TTAATGTCAG TCTGGTTCCC | 4320 |
| AGCATGGTCA GGCTAGGGAT TTTCCGACCG TGGAGGACTT CCTTGTTAAG AATATCCATA | 4380 |
| CCTGCTCGGC CCATTTCTTC AGTATAAACT GTAATACTAG AGAGGGGAGG ATAGACCTGT | 4440 |
| TTGGTCAGAC TAGTGTCTGT AAAGGAAATG AGGCTGACGC GATCTGGCAG GCTGATTCCA | 4500 |
| GCTTCTTGGA GGGCACGGAG GGCACCGATA GCTAACTAT CGCTGGCTGC GAAAAATGCT | 4560 |
| GGCGGAAGTT GGTCTCCCAA GCTCTGAATG GCCTCCTTCA TTAAGTCATA GCCAGACTGG | 4620 |
| GCAGTAAATC TTCTTTGAAA GACCAGTTCA TCATGATAGA TTCCCCTCGC TTGACTATAG | 4680 |
| TTTTTTGAAGT TTTCTAGACG CTGTCTCTGA ATGATTTCTT CTGCTCTGT TGTCTCTCA | 4740 |
| AGGCCTGTGA GAATCCCGAT ACGGTCCATT CCTTGACTGA GGAAATAATC GACAACCTGT | 4800 |
| TTCATAGCAG TGTAAAAATC CGTGATAATA CAGGTATGTC CCAGGGAAAG TGTATCGCTG | 4860 |
| TCTAGAAATA CAAGAGGCTT TTGGTATTCT TCAAAGGCAG AAATCTGAGC TCGACTAAAC | 4920 |
| TTTCCGATGC AGAGAATCCC AATCACTTCC TCGCTTAGGG TAAAAGGGTG GTCATTAAAA | 4980 |
| TAGCGCAAGA TATCATAGTC CAACTCTTGG GCTCTTTTTT CTATTCTTAG GCGAATCTGG | 5040 |
| TAGTAGTAGA GGTCTGCCAG CTCCCCTTGT TCGCTGACCC ATTGGATAAT GGCAATCTTT | 5100 |
| TGCTTGGGTT TGTGGGACTC GCCTGTCTTG AGGTGCTTGG TGTAGCCCAG CTCTTCAGCA | 5160 |
| ACGGTTAAAA TACGGTGTCT GGTTCCTTCT GTAACAGATA GGCTCTGGTC GCGGTTGAGG | 5220 |
| ACGCGGGATA CGGTCGCGAT AGAGACAGAG GCTAGCTGTG CAATGTCTTT TAAGGTAGCC | 5280 |
| ATAAATCCTC CTTGATTAGG TTAGTATATC ATGTTTTTCT TCTTTTACT GATATTTTAC | 5340 |
| TAAAAATTTA GTAAAAAGGA TTGACCTTGG AAAATTCCTT GGATATAATA GAAAGAAAAC | 5400 |
| GATTACACGT TAAGATGGCT TAACGGACAG TCAAAGGAGA ATTATATATG CACAACATCT | 5460 |

621

| | |
|---|------|
| TACTACTGAA GCCCTTCGCA AAGACTTTCT TGCTGTTTTT GGTCAAGAAG CAGATCAAAC | 5520 |
| CTTCTTTTCA CCAGGCCGCA TTAATTTGAT TGGTGAACAC ACAGACTACA ACGGTGGGCA | 5580 |
| CGTTTTTCCT GCTGCTATTT CCTTGGGAAC TTACGGTGCA GCTCGTAAGC GTGACGACCA | 5640 |
| AGTCTTGCGT TTCTACTCAG CTAACCTTGA GGACAAGGGC ATTATCGAAG TGCCTCTCGC | 5700 |
| TGACCTCAAG TTGAAAAAG AGCACAACCTG GACCAATTAT CCAAAAGGTG TCCTTCATTT | 5760 |
| CTTGCAAGAA GCTGGGCACG TGATTGACAA AGGTTTTGAT TTTTATGTTT ATGGAATAT | 5820 |
| TCCAAATGGT GCTGGCTTGT CTTCTTCTGC ATCCTTGGA CTCTTGACAG GAGTCGTGGC | 5880 |
| TGAGCATCTC TTTGATTAA AATTAGAGCG TCTCGATTG GTTAAAATCG GCAAACAAAC | 5940 |
| AGAAAACAAC TTTATCGGAG TAACTCTGG CATTATGGAC CAGTTTGCTA TTGGTATGGG | 6000 |
| GGCAGACCAA CGTGCTATTT ACCTAGATAC TAATACTTTA GAATACGACT TGGTGCCACT | 6060 |
| TGATTGAAG GACAATGTCG TTGTTATCAT GAACACCAAC AAACGCCGTG AATTGGCGGA | 6120 |
| CTCTAAATAC AATGAACGTC GTGCTGAGTG TGAAAAAGCA GTGGAAGAAT TGCAAGTTTC | 6180 |
| CTTGGATATT CAGACTCTGG GTGAATTGGA CGAGTGGGCC GTTGACCAAT ATAGCTATCT | 6240 |
| GATTAAAGAT GAAAATCGTT TGAAACGTGC TCGCCATGCT GTGCTTGAAA ACCAACGTAC | 6300 |
| CCTCAAAGCT CAAGTAGCAC TCCAAGCAGG AGATTTGGAA ACATTTGGAC GCTTGATGAA | 6360 |
| TGCGTCACAC GTTCTCTGG AGCATGATTA TGAAGTAACT GGTTTGGAAT TGGATACCCCT | 6420 |
| TGTTACACAC GCTTGGGCAC AAGAAGGAGT TCTCGGTGCT CGTATGACAG GGGCTGGTTT | 6480 |
| TGGTGGCTGT GCCATTGCCT TGGTTCAAAA AGATACTGTT GAGGCCTTTA AGGAAGCTGT | 6540 |
| AGGCAAACAC TACGAGGAAG TAGTTGGATA CGCTCCAAGC TTCTATATCG CTGAAGTTGC | 6600 |
| AGGTGGCACT CGCGTCTTG ACTAGTCAAA AGGAGGCTCT ATAGTGACCT TAGTAAATAA | 6660 |
| ATTTGTAACA CATGTCATTT CTGAAAGCTC ATTTGAGGAA ATGGATCGAA TCTATCTGAC | 6720 |
| CAATCGTGTT TTGGCACGAG TGGGAGAAGG TGTTTTGGAA GTTGAGACCA ATCTGGATAA | 6780 |
| ATTGATTGAC CTCAAGGACC AGCTGGTTGA AGAAGCCGTT CGATTAGAGA CGATTGAGGA | 6840 |
| TAGTCAGACT GCGCGTGAAA TCCTTGGTGC TGAACGTATG GATTTGGTGA CTCCTTGTC | 6900 |
| AAGTCAGGTC AATCGTGATT TTTGGGCAAC CTACGCCAC TCTCCAGAAC AAGCGATAGA | 6960 |
| GGATTTTAC CAACTCAGTC AGAAAAATGA CTACATCAAA CTCAAGGCCA TTGCTAGAAA | 7020 |
| TATCGCTTAT CGTGTTCAT CTGACTACGG AGAAGTTGAA ATTACCATCA ATCTCTCTAA | 7080 |
| GCCTGAAAAA GATCCCAAAG AGATTGTGGC AGCCAAGTTG GTGCAAGCTA GTAATTATCC | 7140 |
| TCAGTGTGAG CTTTGTCTAG AGAATGAGGG CTACCATGGT CGAGTTAACC ACCCAGCTCG | 7200 |

622

| | |
|--|------|
| TAGCAATCAC CGTATTATCC GTTTTGAAAT GGTTGGTCAG GAATGGGGTT TCCAGTATTC | 7260 |
| GCCCTATGCT TACTTTAATG AGCATTTGTAT CTTTTTAGAT GGCCAGCATC GTCCCATGGC | 7320 |
| CATTAGTCGT CAGAGTTTGT AACGTCTGTT GGCTATCGTA GACCAGTTTC CAGGATATTT | 7380 |
| TGCTGGATCT AATGCCGACC TGCCGATTGT GGGGGGCTCT ATTCTAACTC ATGATCATTA | 7440 |
| TCAGGGAGGC CGTCACGTAT TTCCTATGGA ATGGGCTCCC TTGCAAAAGG CCTTCCGATT | 7500 |
| TGCTGGTTTT GAGCAGGTCA AGGCTGGAAT TGTCAAGTGG CCCATGTCTG TCCTACGTTT | 7560 |
| GACTTCGGAT TCCAAAGAGG ATTTGATCAA TTTGGCTGAT AAGATTTTGC AGGAATGGCG | 7620 |
| CCAGTATTCA GATCCTGCAG TGCAGATTTT GGCAGAGACA GACAGGACAC CGCATCACAC | 7680 |
| TATCACACCC ATTGCCCCGA AACGCGATGG ACAGTTTGAG TTGGACTTGG TCTTGCGAGA | 7740 |
| CAATCAGACT TCAGCAGAGT ATCCTGATGG TATCTATCAT CCCCACAAGG ATGTCCAACA | 7800 |
| TATCAAGAAG GAAAATATCG GCTTGATTGA GGTATGGGC TTGGCAATCT TGCCACCACG | 7860 |
| TCTGAAAGAA GAAGTGGAGC AAGTCGCTAG CTATCTTGTA GGAGAAGCTG TTACAGTTGC | 7920 |
| CGATTATCAT CAGGAGTGGG CAGACCAACT CAAATCCCAA CATCCAGACT AACGGATAAA | 7980 |
| GAAAAAGCCC TTGCAATCGT CAAGGACTCT GTGGGTGCTA TCTTTGCGCG TGTACTTGAG | 8040 |
| GATGCAGGAG TCTACAAGCA GACAGAACAA GGCAGACAG CCTTTATGCG CTTTGTGGAA | 8100 |
| CAGGTCCGAA TTTTACTAGA CTAGGAGCTT TCTCGG | 8136 |

(2) INFORMATION FOR SEQ ID NO: 76:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10011 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76:

| | |
|---|-----|
| CCCATAGTGA AGAGTGGCCA TAAGAAGGTC TTCTAGGCTT AATTAGGTT TTCGTCCACC | 60 |
| TTTTGCGTGT TTAAGTTGAT AAGCTGTTTT TAACACAGCT GAACATCTCT TCAAAAGTCG | 120 |
| TGCGCTGAAC ACCAACAAGA CATTAAATC GTGTATCAGT TAGTTGTTTA CTGCTTCAT | 180 |
| CATTCATAGA ACTACTATAC CATGTTTTGT TTCGCAGGAA GTCTAATATT GTCAAATACT | 240 |
| GGAACGCTCA TTGCTGGGAT ACGGAATAAG ATGGGCCAG CTTCGATAAC TGGGATACCT | 300 |
| GGTTCAAAAC CAAGGTCTGT TGCAGCGATT GGTGTAAAGA TATCGTAACC TTTCATAAGG | 360 |
| TCTTCGTTTA CATCTTTCAC CATAACTGCA TCACAGTGAA CATCGTAACC ACGGTTTGAA | 420 |
| AGTTCTTCTT CTAGAGCACT TTTAATTTGG TGAATTGAGT TAACACCTGC ACCGCAGGCA | 480 |

623

| | |
|--|------|
| GCAAGAATTT TAATCATTTG GATTTCCTCC GATTTTATTT TTTAATAGAC AAGATTAAGC | 540 |
| GGTTGCTTCA GCAATGTAAG CATAAAGGGC TTCTGGTTCA GAAATTTTGG ATAGGCTTTC | 600 |
| AAGATGACCA TTTCTGTGA AGAAGTCCAT TAACTGAGCA AGAATGTTCG TTTGACTTGA | 660 |
| ACTTGAATTA TTGATGATAA AGAAGAGCAA GGATACTTCT ACTTCCTTAC CTGGCGCAAT | 720 |
| CATATTATGG AAAGTCACCG GTTCTCTTAA TCGAACAACC ACCACTTTCT CAGCTAGATT | 780 |
| ATGAACAATA TCTGTGTGAG GAATCATTAC ATTTGCAAGT CCTTTCCTAG AAATCCATA | 840 |
| TATAAACCAG TTGGAAATGA CTTTTCACGC GTGATCAAGG CTTACAGATA AGTTGGAGTG | 900 |
| ACAATTTCTC GTTCTTCCAA CAAGCTTGCT ACCTGATCAA AAAGTTATTC TTGATTATCC | 960 |
| GCTTCTAAGC AAAACACAAG GTTTTTGTCA AAGAAATAAT CTAATACCAT AAGGTTTTCC | 1020 |
| CTTCTTTCCA TTAACCTTAT GCTATAAGTA TAACACTATA TGAAATCGTT GTTAATTACT | 1080 |
| TTCTATTCTT TTTTGTCTCT TTTTTATAT TTTTGTTTG TTTATAGTTT GTTATATAAA | 1140 |
| AATAAACACA CAAACAAATA CTCCAAGCAT TTTTCTGTTC TAATACTCAA TGAAATCAA | 1200 |
| AGAGCAAACCT AGGAAGCTAG CCGCAGTTGT TCAAAACACA GTTTTGAGGT TGTAGATGAA | 1260 |
| ACTGACGAAG TCACTCAAAA CATGGTTTGG AGGTTGTAGA TGAAACTGAC GAAGCAACAG | 1320 |
| CCATACATAC GGTAAGGCGA CGCTGACGTG GTTTGAAGAG ATTTTCGAAG AGTATAAAAA | 1380 |
| CTAAAAAAGC AGACCATCTA AGCCTGCTTT ACTATTGATT CTTATATAAA TTTCTGTGA | 1440 |
| ACAAGGAAAG GCATTTCTGA TAACTTATTC TTCATCCATA CTCAAGACGC TGAGGAAGGC | 1500 |
| TTCTTGCGGA ACTTCAACTG ATCCGATGGA TTTTCATGCGT TTCTTACCAG CTTTTTGT | 1560 |
| TTCAAGGAGT TTACGCTTAC GAGAAACGTC ACCACCATAA CATTTAGCAA GTACGTTCTT | 1620 |
| ACGAAGGGCC TTGATATCAG TACGAGCGAC AATCTTGTGT CCAATAGCCG CTTGGATTGG | 1680 |
| AACCTCAAAT TGTTGGCGAG GGATGATTTT CTTGAGTTTA TCAACGATGA GTTTCCACG | 1740 |
| TTCTAGGCA AAGTCCTTGT GAACGATAAA GCTGAGGGCA TCCACCTTAT CTCCATTGAG | 1800 |
| AAGAATATCC ATTTTCACCA GCTTAGATGG GCGATATTCT GACAATTCGT AGTCAAAGCT | 1860 |
| TGCATAACCA CGTGTGCAAG ACTTAAGTTT ATCAAAGAAG TCAAAGACAA TTTCAGCAAG | 1920 |
| AGGAATTTGA TAGATAACAT TGACACGGTT ATCATCAATA TAGTCCATAG TCACAAAGTC | 1980 |
| CCCACGCTTA CGCTGAGCTA GCTCCATTAC TGCTCCGACG AACTCCTGTG GTACCATGAT | 2040 |
| TTGCGCCTTG ACATAAGGCT CTTCAATGGT CGCAATCTTA GTTGGGTCTG GAAACTCAGA | 2100 |
| TGGGTTAGAC ACATCCATAG ACTCACCCTC GGTCAAATTA ACTTTGTAAA TAACAGACCG | 2160 |
| AGCTGTCTATG ATGAGGTCAA TATTGAACTC ACGCTCTAAA CGTTCCTGGA TAACATCCAT | 2220 |

624

| | |
|---|------|
| ATGGAGAAGT CCAAGAAATC CACAACGGAA ACCAAATCCA AGTGCCTGAG ATGTTTCTGG | 2280 |
| TTCAAACCTGA AGACTAGCAT CATTCAGTTG CAATTTTTCAG AGCGCTTCAC GCAGGTCATT | 2340 |
| GTACTTGTTT GATTTCGATTG GGTAGAGACC CGCAAAGACC ATAGGATTCA TCTGCTTATA | 2400 |
| ACCATGTAAT GGTTCGCGG CAGGATTGGT TGCCAAGGTA ACGGTATCAC CCACACGAGT | 2460 |
| ATCCTGAACC GTCTTGATAG ACGCCGCAAT GTAACCAACA TCACCAGTCG CAAGGAAATC | 2520 |
| ACGACCAACC GCTTTTGGTG TAAAAATACC GACTTCGGCC ACATCAAAGG TCTTACTATT | 2580 |
| GCTCATGAGC TGAATCTTAT CACCAGGTTT GACCACTCCG TCCATGACAC GCACCTGGAG | 2640 |
| GATAACCCCA CGGTAAGCAT CGTAAACAGA GTCGAAAATC AAGGCCTTAA GTGGCGCCGT | 2700 |
| CACATCACCC GTTGGTGCTG GTACTTTTTC TACAATTTGC TCGAGGATTT CTTCAATCCC | 2760 |
| AATACCAGCC TTGGCAGAAG CCAAACTGTC TTCACTGGCA TCCAAACCAA TCACATCTTC | 2820 |
| AATCTCTGTA CGCACGCGCT CCGGATCTGC AGCCGGCAGG TCAATTTTAT TAATGATAGG | 2880 |
| CATGATTTCC AAATCATTAT CCAAAGCCAG ATAAACGTTG GCAAGAGTTT GAGCCTCAAT | 2940 |
| TCCTTGAGCC GCATCGACCA CCAAAATAGC ACCCTCACAG GCAGCTAGCG AACGTGAAAC | 3000 |
| TTTCATAGGTA AAGTCAACGT GCCCTGGTGT GTCAATCAAG TGGAAAATAT AAGTTTCCCC | 3060 |
| ATCTTTTGCA GTGTAATTCA ACTCGATGGC ATTCAACTTA ATAGTAATTC CACGTTCCCG | 3120 |
| CTCTAGCTCC ATGCTATCCA AAAGCTGGGC CTGCATTTCG CGACTTGAAA CCGTCTCTGT | 3180 |
| TTTTTCCAAA ATGCGGTCTG CTAGAGTTGA TTTTCCGTGG TCAATATGGG CGATAATAGA | 3240 |
| GAAGTTACGG ATCTTCTCCT GTCGTTTTTT CAATCTTCTT AAGTTCATGA TTCTCTCTCT | 3300 |
| TTCAGGGTAT CTATTTATTA TAAATTGTTT TTGATATTTT GACAAGACCA TACCCTGCTA | 3360 |
| GGAGTACTAA TCTTCAGCGA CAAAGCCGTC ATTTTCGATA AAGTGCTGTT CTGTCATTCC | 3420 |
| TTGGTCTGTA AAGACAATCC CGTGAAGGAC ACCACCATAA ACAGCTCCTC CATCCATTCC | 3480 |
| AATCTTGCCA TCTTCTGTAG TCCAAAGCTC AGATGTACCG CGTTCTTGCT GTAACAAACC | 3540 |
| ATAGACCGGT GTATGACCGA AGACAATGGT TTTTCCAGTA TGATTTCAG CTCCGTGGAA | 3600 |
| TGGTTTTCTA AGCCATACTT TTTTATAATC TGTGTTTCA TGCCAGTCGT CCAAGGTCAA | 3660 |
| ATCAATACCT GCGTGAACAA AGATATACTT GTCTGTCTCT ACTACAAATG GCATTTGACG | 3720 |
| AATGAATTCTG ACCAAGTCTG CCGCTTCAGC GgCAACCCGC TTGGCATCTT CTACTCCATC | 3780 |
| AACTGGTGCA TCCAAGGGAC GACCTAGGAT AGAGTTAATG GTTGTATCTC CACCATTGCG | 3840 |
| ACTATAATGG TCATAACTTT CTTCTGGGTC ATCTAGCCAA GTCAAAAACA TATACTCGTG | 3900 |
| GTTTCCGGAC AAACAGATAG CCCCTTGATT GTCCACCAAG TCCTTGACCA TTTCAAGAAC | 3960 |
| ACGGTGACTA TCCTCACCTC TGTCAATCAA ATCACCTAGA AAGAGCAACT GGGGCTGACC | 4020 |

625

| | |
|--|------|
| ATCCCAGGTT TTGAGAAGGT CTTCAGCAT CCCAGCTTTT CCGTGAACAT CTCCAATTAC | 4080 |
| ATAATAATCT GTCATCTTAT TTCTCCCTGT TTCTCAACAA TTCTCTTGCT TGCGTCAGGG | 4140 |
| CTGCTTCTGT CACATCATCA CCTGCCAACA TCTTGGCAAC TTCTTCCACT CGCTCTTCGA | 4200 |
| CCGTCAAGAG ACGAACAGTC GAAACCGTTG AATGGTCATT ACTAATCTTC TCAATAAAGA | 4260 |
| ATTGATAATC TGCAATCGCA ATTACTTGTG GCAAAATGGGA GATAGCCAAA ACCTGACCAT | 4320 |
| GCTGACCAAT TTTATGAATT TTCTGAGCAA TAGCTTGAGC AACACGACCT GAAACTCCCG | 4380 |
| TATCCACCTC ATCAAAGACA ATGCTAGTCT TGCCCTTCTTT ACGTGAAAAG GCAGACTTAA | 4440 |
| TGGCTAACAT GAGACGAGAT AATTCCCTC CAGAAGCAAC CTTAACCAAG GGTTTAAAGT | 4500 |
| CTTCTCCAGG GTTGGTTGAA ATATAAACT CAACCATTTT ATTTCCCTCA CGACTGAATT | 4560 |
| TTCCCTTACT AAAACGAACC TGAACTGGG CTTTTTCCAT ATAAAGATCT TGCAGTTCTT | 4620 |
| GTTTAATCTC AGCTTCGAGT TGCTGAGCCA AATTATGACG AGCAGAAGCA AGTTGACCTG | 4680 |
| CCAAATTGAC AAGATTGACT TCCAACTTCT TAAGCTCTGC TTCCATGTCC TCAGACGAAA | 4740 |
| GATTATTGCC TGTCAAGAGA TTGTATTCTT CCGTAATCTT GGCAAAATAA AGCAAAACAT | 4800 |
| CATCAACAGT CCCACCATAC TTACGAGTAA TAGTATGAAG GAGGTCCAAA CGATTCTCAA | 4860 |
| CCTGCATCAG GCGATTGCCA TCAAAATCAA GGTCTCAAT GATAGCTTCC AAACGTTTGC | 4920 |
| TAATGTCTTC TAAAACATAG TAGGTCTCAG ACAGATAGCT TGAAATTTCA CGGTATTCAG | 4980 |
| GATCATACTC TTCGACACTT TCCATGTCTT TCATAGCTGA ACGAACATTG GCCAGACTTG | 5040 |
| AAAAATCTTC ATGTGCTAAC ATACTGTAGG CATTTGGTCAG TGTATCCGCA ATATTTTGT | 5100 |
| GGTTGAGGAG TTTATCTCGC TCTTGATTGA GAGCCAAGTC TTCTCCAGCC TGCAAGTTTG | 5160 |
| CTGCCTCAAT CTCTGCCATT TGAAATTCCA ACATTTGAT ACGTGCCTTG TGTTCCTGTT | 5220 |
| GGTTTTCTT GACTTCCAGA ACCTGCTTGC GCATTTTCCG ATAGGCATCA AAACCTGTTT | 5280 |
| GATAGGTTTC TTTCAGTCC CAAAAGCGG CATCACCAAA TTCATCCAAC ATCTGGATAT | 5340 |
| GCAGTTGGGG ACGCATTAAC TCCTCATGGT CATGCTGACC ATGAATATCT ACAAGATGTT | 5400 |
| GCCCAATAGC TCGCAAAACA GACAGATTAA CCATCTGACC ATTTACACGG CTGATACTAC | 5460 |
| GACCATTTTG CAAGATTTC CGACGGATGA TAATTTATC ACCTAATTCT AAACCTTGCT | 5520 |
| CATCAAAAAT TTCTGTAAA AGACGACTAT TCTCAACTGA GAAAAGCCCC TCAATCTCTG | 5580 |
| CCTTTGGTGC ACCATGACGA ATAACATCTG TCGTCGCACG AGCTCCCAAC ATCATATTCA | 5640 |
| TGGCATCAAT GATAATCGAC TTCCCTGCAC CCGTTTCACC AGTCAGGACA GTCATCCCTT | 5700 |
| TTTCAAAATT GAGGGAAATA GCCTCAATAA TGGCAAAGTT TTTTATCGAA ATTTCAAGTA | 5760 |

626

| | | |
|---|-----------------------|------|
| ACATATAGAC CTACCAATTT TTTACTTGTT CAAAGATTTT | CTCTGCTAGA CTTCCACTTC | 5820 |
| TGGCAATGAC TAAATCGAG CTATCATCAG TCAAACAGCT | AAAAATCTTG TCTGCAAAAG | 5880 |
| TCTCGATTAA CTGAGCTTTT ACAAAGCCG TATTTCTGG | AATAACTTGG AGATTGATCA | 5940 |
| TCTTATCCAT CAATTCAGCC GATTCGATAT TGTCTTCAGC | CAGTTGCAGA CTTTTTACGA | 6000 |
| TTGATTTTGG CAATTCGTAG ACATAGGTGT TGTCTCTCAA | AGGAATTTTG ACAATACCTA | 6060 |
| ACTCTTTGAT ATCTCGGGAT ACCGTCGCCT GAGTGGCAGT | GATACCTGCT TCTTTCAAAT | 6120 |
| GTCTACAAT TTCTCTTGC GTGCCGATTT GATAATCTGT | CACCAATCTT CTAATTTTMT | 6180 |
| CAAGTCTCTC TTTTATATC ATTTTAAAT TGAATATGCG | CCCTCTCTAC TGCTTCTTTA | 6240 |
| ATCTCAGCAA GAATCTGATT GCTTGCTGAC TTTTCTTTT | TCAAATACGC TAAAAATTCA | 6300 |
| ATATTTCCAT GTCCACCTTG GATGGGAGAA AAGTCCAAGC | CAAGGACTGA AAAACCTACC | 6360 |
| TCTACTGCCA TAGCTGTTAC AGATTCAAGG ACATTCTGAT | GAACCTTAGC ATCTCGAATA | 6420 |
| ATTCCATTTT TCCCAATCTG CTCACGTCCT GCCTCAAAC | TGAGTTTGAC AAGTGCTACC | 6480 |
| ACCTGACCTT GATCAGCCAA GACACGGTGC AAGGCTGGCA | AAATCAGACT AAGGGAAATG | 6540 |
| AAACTCACAT CAATACTGGC AAAGCTCGGC TCCTGCTCGA | AATCAGTCTT TTCAGCATAG | 6600 |
| CGGAAATTGA ACTGCTCCAT GCTGACAACT CGTGGGTCTT | GGCGTAATTT CCAAGCCAAC | 6660 |
| TGATTGGTAC CAACATCGAC TGCAAAGACC AACTTGGCAC | TATTCTGTAG CATGACATCG | 6720 |
| GTAAACCTC CAGTAGAGGC CCCGATATCA ATCGTAGTCG | CGCCATCCAC CGACAAATCA | 6780 |
| AAGACCTGCA AGGCCTTTTC CAGTTTCAA CCACCACGGC | TGACATACTT GAGTTTCTCC | 6840 |
| CCCTTGAGTT TTAATTCGGT GTCATCTGGA ATTTTCTCTC | CTGGCTTGTC AAACCGTTCT | 6900 |
| CCATTAAGGA CTGCTACGAC TAGGCCAGCC ATCACACCTC | GCTTGGCCTG CTCTCTCGTT | 6960 |
| TCAAACAACC CCTGTTTATA AGCTAGTACA TCCACTCTTT | CCTTAGCCAT TGATTCTCAA | 7020 |
| ACTTTCTACT ACATTACAA TCGATTCTGT TTCAAAGGGA | AGCTGCTGGG CAATTTCTTC | 7080 |
| TAATTTTCA TTAGCTTGAT CCAGGGTTTG GTTACAAAAG | GCAATGGACT CTTCCAAGCC | 7140 |
| CAACAGGGCA GGATAGGTTG ATTTTCTGTC CTGCAGATCC | TTTGAGGTG TCTTGCCGAT | 7200 |
| TTCTCAAAA CTAGCTGTCA CATCCAGTAC ATCATCTCTG | ACTTGAAAAG CAAGTCCAAT | 7260 |
| CAATTACCCC ACAGTTTCA GCTTCACCTG CATTCAGGT | GACAATTCAG CTATAATAGC | 7320 |
| TGCCGCTTGG AAGGATAGG CTAGTAACTT CCCAGTCTTA | TTGGCATGAA TAGTCTGAAG | 7380 |
| TTCTTCCAAA GACAAGTCT GGTGTTTCGCC CTCCATATCC | AAAACTTGCC CTGCTACCAT | 7440 |
| ACCCAGACTA CCTGAAGCAA GGGATAAGTT GGCAATCAAG | TCCACCTTAA TCTGACTTGG | 7500 |
| CAAATCTGCC TGCGAATCA AGGCATATGA GTCTAAGAAT | AAGGCATCTC CAGCCAAAAT | 7560 |

627

| | |
|--|------|
| GGCCATAGCT TCACCGAATT TCTTGTGATT GGTAAACCGC CCTCTTCGAT AATCGTCATC | 7620 |
| ATCCATAGCA GGAAGGTCAT CGTGAATCAA GCTCCCTGTA TGAATCATCT CTAAGGCAGT | 7680 |
| AGCTACCTGC GCGTGAGCAG GTTTGATGGT AACCTGCAAG GCTTCCAGAA CTTCTAACAA | 7740 |
| GAGAAAAGGC CGAATACGCT TGCCACCAGC ATGAATAGAA TAGAGAACAG ACTCCCGTAA | 7800 |
| ACTAGAGGCA AACTGCTGGT CTCCATAAAA ATCTTCCAAA GCCGACTCGA CAAGAGCTAA | 7860 |
| TTTTTCTTGC TTTTTCATTC AAAATCACTT TCTGTTCCGT CTTCTTGCAT GACCTTGACC | 7920 |
| AAGGTCTTTT CAGCCTTGTC CAGCGTAGCT TGGAGCTCTT TTGACAAGAC CATGCCCTTT | 7980 |
| TGAAAGGCAG TAATCGCATC TTCCAGAGCA ATTTCACCAT TTTCCAAACT TTGGACAATG | 8040 |
| GTTTCCAGTT CTGCTAGATT TTCCCTCAAAT TTCTTTTGGT TTGACATCTT TAACCTCTAA | 8100 |
| TTCTACTTGA CCATCTCGCA TCAAAAGCGT TACTTGGTCT TTTTCTTCA AACTCTCAAC | 8160 |
| CGAATCTACA ACGGACTCTT CTTTTTTGAC AATAGCATAA CCACGCGCCA CGATTGCGCT | 8220 |
| AGTATCCAAC ATGAGCAAAG CTTCCGAAAG TCGCTTGGCC TCAGCAACCT TGGCGTCATA | 8280 |
| AACTAACGCC ATTTGGCTAC CTAAGAGCTT GTCCAACGTG CCTAAACGGT CTTGATAGCG | 8340 |
| TTGGATTTTG GTAACAGGTG ATAATTGTAC TAATTGATGA GTTCTTGCTT GAACTAATTG | 8400 |
| TTTGTTATCA GAAATCCGAG TTCGCAAAC TTTGTTTCAA CGCAGTTGCA GTTGGTCCAA | 8460 |
| GCGTTGCAAA TAACCGTCAT ACAAGCGCTC AGGTTGTCTA AAGATAACAG ACTGACTGCA | 8520 |
| TTTTTTTCAA GCCTCTTGTT TCTTAGATAG AACATTTGCG ACTGCCGTTA CCATCCGTTT | 8580 |
| TTCTTGATTT TGCAAAATGAG CTAATACATC CAACTTGGTC ACAGGTGTTG CCAGTTCAGC | 8640 |
| CGCCGCTGTT GGCCTTGACG CGCGTCGATC TGCCACAAAA TCTGCCAAGG TCACATCCGT | 8700 |
| CTCATGCCCC AACTAGAGA TAACTGGCAA ACGAGATTCA AAAATAGCTC GTACCACAAT | 8760 |
| TTCTTCGTTA AAGGCCCAGA GATCCTCAAT AGAACCACCT CCACGACCAA TAATGAGCAA | 8820 |
| ATCCAAATCG TCCCGTTGAT TAGCACGCGC AATATTTCTA GCAATTTCTT CCGCAGCCCC | 8880 |
| TTACACCTGA ACCTTGGTCG GATAAAGAAG GATGTCAACA CCTGGGAATC GCCTGCTGAC | 8940 |
| GGTCGTGATA ATATCTCGAA TAACGGCTCC ACTACGGCTG GTTACTACAC CAATTCCTT | 9000 |
| AGAAAATTGG GGCAGAGCTT GCTTGAAGCG TTCTTGAAAC AGGCCTTCTT CTGTCAATTT | 9060 |
| TTTCTTAAGT TGTTCAAACT GAATCGCAAG CGCCCCAACC CCATCAGGCT CAGCTTTTTC | 9120 |
| AATGATGATG GAGTAGCTAC CACTTGGTTC ATAGACCTGT ACACGCCCAA TCACATTGAT | 9180 |
| CTTCATTCCT TCTTCCAGGT CAAACCCTAA TTTCTGATAA ATCCCAGACC AGATGGTCGC | 9240 |
| TTGAATAACT GCATGGTCAT CCTTTAGGGA GAAATATTGG TGAGTAGGTC GTTTACGAAA | 9300 |

628

| | |
|--|-------|
| GTTGGAAACT TGACCAGTTA AATAGACCCG TTCCAAGTAT GGGTCTTTAT CGAATTTTCAT | 9360 |
| TTTCAGATAC TTGGTCAAAG TTGTTACCGA TAAATACTTT TCCATCTCCA CCTACTATTC | 9420 |
| ATTTACTTGC TCTTTCATGG GTATTATTAT ACCAAAAATA TGCCTAAAAA TCTCCATTTA | 9480 |
| TGTACCATTA TGAGGGAAAA ATAGAAAAAG GAGGCAAGGC CTCCACATGT GATTATTTGC | 9540 |
| TGTTTCGAGC TTCTTCCAAA ATCTTTGCAA TCTTGGTCGT CAACAGGTCG ATAGCCACGG | 9600 |
| TATTGCTAAC CCCTTCAGGA ATGACGATAT CAGCATAACG CTTAGTTGAC TCGATAAACT | 9660 |
| GGTGGTACAT TGGTTTGACC ACACCTAAGT ACTGGTTAAT AACGCTATCA AGGCTACGGC | 9720 |
| CACGCTCCTC CATATCACGC TTGATACGAC GAATAATGCG CACATCGTCA TCCGTATCCA | 9780 |
| CAAAAATCTT GATATCCATC AAATCGCGCA GACGCTTGTC CTCCAAGACC AAAATACCCT | 9840 |
| CAACGATAAA GACATCTTGA GGTTCCTGAC GATAGGTCTT GCTACTCCGT GTATGCTCTG | 9900 |
| TATAGTCGTA GGTCCGGATG TCCACCGGAC GCCCTGCCAA CAATTCCTTA ATCTGCTCGA | 9960 |
| TCATCAAGTC TGTATCAAAG GCAAAAGGAT GGTCATAGTT GGTTTTGACG G | 10011 |

(2) INFORMATION FOR SEQ ID NO: 77:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 5365 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 77:

| | |
|---|-----|
| CGTGTGGTCT TAAAAATAGA AGACAAAGAA CAACTGTTG GAGGCTTTGT CCTTGCAGGC | 60 |
| TCAGCCCAAG AAAAAACCAA AACAGCTCAA GTGTGGCTA CTGGACAAGG TGTTCTGACC | 120 |
| TTGAACGGTG ACTTGGTTGC TCCAAGTGTT AAAACTGGAG ATCGTGTCTT AGTTGAAGCC | 180 |
| CACGCAGGTC TTGATGTCAA AGATGGCGAT GAAAAGTACA TCATCGTAGG CGAcTAACAT | 240 |
| TTTGGCAATC ATTGAGGAAT AGAAGGAGAA AGTAAGTATG TCAAAAGAAA TTAAATTTTC | 300 |
| ATCAGATGCC CGTTCAGCCA TGGTTCGTGG TGTCGATATC CTTGCAGACA CTGTTAAAGT | 360 |
| AACCTTGGGA CCAAAAGGTC GCAATGTCGT TCTTGAAAAG TCATTCCGTT CACCCTTGAT | 420 |
| TACCAATGAC GGTGTGACCA TTGCCAAAGA AATCGAATTG GAAGACCATT TTGAAAATAT | 480 |
| GGGTGCTAAG TTAGTATCAG AAGTAGCTTC TAAAACCAAT GATATCGCAG GTGACGGAAC | 540 |
| TACGACTGCA ACAGTCTTGA CCCAAGCTAT CGTCCGTGAA GGAATCAAAA ACGTCACAGC | 600 |
| AGGTGCAAAT CCAATCGGTA TTCGTCGTGG GATTGAAACA GCAGTTGCCG CAGCAGTTGA | 660 |
| AGCTTTGAAA AACAACGCCA TCCCTGTTGC CAATAAAGAA GCTATCGCTC AAGTTGCAGC | 720 |

629

| | | | | | | |
|------------|-------------|------------|------------|------------|------------|------|
| CGTATCTTCT | CGTTCTGAAA | AAGTTGGTGA | GTACATCTCT | GAAGCAATGG | AAAAAGTTGG | 780 |
| CAAAGACGGT | GTCATCACCA | TCGAAGAGTC | ACGTGGTATG | GAAACAGAGC | TTGAAGTCGT | 840 |
| AGAAGGAATG | CAGTTTGACC | GTGGTTACCT | TTACAGTAC | ATGGTGACAG | ATAGCGAAAA | 900 |
| AATGGTGGCT | GACCTTGAAA | ATCCGTACAT | TTTGATTACA | GACAAGAAAA | TTTCCAATAT | 960 |
| CCAAGAAATC | TTGCCACTTT | TGGAAAGCAT | TCTCCAAAGC | AATCGTCCAC | TCTTGATTAT | 1020 |
| TGCGGATGAT | GTGGATGGCG | AGGCTCTTCC | AACTCTTGTT | TTGAACAAGA | TTCGTGGAAC | 1080 |
| CTTCAACGTA | GTAGCAGTCA | AGGCACCTGG | TTTTGGTGAC | CGTCGCAAAG | CCATGCTTGA | 1140 |
| AGATATCGCC | ATCTTAACAG | GCGGAACAGT | TATCACAGAA | GACCTTGGTC | TTGAGTTGAA | 1200 |
| AGATGCGACA | ATTGAAGCTC | TTGGTCAAGC | AGCGAGAGTG | ACCGTGGACA | AAGATAGCAC | 1260 |
| GGTTATTGTA | GAAGGTGCAG | GAAATCCTGA | AGCGATTCTT | CACCGTGTG | CGGTTATCAA | 1320 |
| GTCTCAAATC | GAAACTACAA | CTTCTGAATT | TGACCGTGAA | AAATTGCAAG | AACGCTTGGC | 1380 |
| CAAATTGTCA | GGTGGTGTAG | CGGTTATTAA | GGTTGGAGCC | GCAACTGAAA | CTGAGTTGAA | 1440 |
| AGAAATGAAA | CTCCGCATTG | AAGATGCCCT | CAACGCTACT | CGTGCAGCTG | TTGAAGAAGG | 1500 |
| TATTGTTGCA | GGTGGTGGAA | CAGCTCTTGC | CAATGTGATT | CCAGCTGTTG | CTACCTTGGA | 1560 |
| ATTGACAGGA | GATGAAGCAA | CAGGACGTAA | TATTGTTCTC | CGTGCTTTGG | AAGAACCCGT | 1620 |
| TCGTCAAATT | GCTCACAATG | CAGGATTTGA | AGGATCTATC | GTTATCGATC | GTTTGAAAAA | 1680 |
| TGCTGAGCTT | GGTATAGGAT | TTAACGCAGC | AACTGGCGAG | TGGGTAAACA | TGATTGATCA | 1740 |
| AGGTATCATT | GATCCAGTTA | AAGTGAGTCG | TTACGCCCTA | CAAAATGCAG | CATCTGTAGC | 1800 |
| CAGCTTGATT | TTGACAACAG | AAGCAGTCGT | AGCCAATAAA | CCAGAACCAG | TAGCCCCAGC | 1860 |
| TCCAGCAATG | GATCCAAGCA | TGATGGGCGG | GATGATGTAA | GCTTTCTATA | GAAAACAACT | 1920 |
| TATAAAAAAC | ACAAAAGGAG | GGAATGACTA | ACCCTTCTTT | TTATAGGCTC | TTTGTCAACT | 1980 |
| GTAGTGGGTT | GAAGTCAGCT | AAGCTCGAGA | AAGGACAAAT | TTCGTCCTTT | CTTTTTTGAT | 2040 |
| GTTCAAAGCG | ATAAAAAATCC | GTTTTTTGAA | GTTTCAAAG | TTTCGAAAAC | CAAAGGCATT | 2100 |
| GCGCTTGATA | AGTTTGATGA | GATTATTGGT | CGCTTCCGGT | TTGGCGTTAG | AATAGTGTAG | 2160 |
| TTGAAGGGCG | TTGATAATCT | TTTCTTTATC | TTTGAGGAAG | GTTTTAAAGA | CAGTCTGAAA | 2220 |
| AATAGGATGA | ACTTGCTTAA | GATTGTCTC | AATAAGTCCG | AAAAATTCT | CCGGTTCCTT | 2280 |
| ATTCTGAAAG | TGAAACAGCA | AGAGTTGATA | GAGCTGATAG | TGATGTTTCA | AGTCTTGTGA | 2340 |
| ATAGCTCAAA | AGCTTGTCTA | AAATCTCTTT | ATTGGTTAAA | TGCATACGAA | AAGTAGGACG | 2400 |
| ATAAAATCGC | TTATCACTCA | GTTTACGGCT | ATCCTGTTGT | ATGAGCTTCC | AGTAGCGCTT | 2460 |

630

| | |
|---|------|
| GATAGCCTTG TATTCATGGG ATTTTCGATC CAATTGGTTC ATAATTTGAA CACGCACACG | 2520 |
| ACTCATAGCA CGGCTAAGAT GTTGACAAAT GTGAAAGCGA TCCAACACGA TTTTAGCATT | 2580 |
| CGGGAGTGAA ACAGTCTGGG AGACTGTTTC AGCCTGAGCC TAGAAATTTG AAAGCGAAGC | 2640 |
| TGTTTAGCCA AGTCATAGTA AGGACTAAAC ATATCCATCG TAATGATTTT CACTTGACAA | 2700 |
| CGAACGGCTC TATCGTAGCG AAGAAAGTGA TTTCGGATGA CAGCTTGTGT TCTGCCTTCA | 2760 |
| AGAACAGTGA TAATATTAAG ATTATCAAAA TCTTGCGCAA TGAAACTCAT CTTTCCCTTA | 2820 |
| GTGAAGGCAT ACTCATCCCA AGACATAATC TTGGAAGCC GAGAAAAATC ATGCTCAAAG | 2880 |
| TGAAAGTCAT TGAGCTTGGC AATGACAGTT GAAGTTGAAA TGGCCAGCTG ATGGGCAATA | 2940 |
| TCAGTCATAG AAATTTTTC AATTAACCTT TGAGCAATCT TTTGGTTGAT GATACGAGGG | 3000 |
| ATTTGGTGAT TTTTCTTTAC CAGGGGAGTC TCAGCAACCA TCATTTTGA ACAGTGATAG | 3060 |
| CACTTGAAAC GACGCTTCT AAGGAGAATT CTAGAAGGCA TACCAGTCGT TTCAAGATAA | 3120 |
| GGAATTTTAG AAGGTTTTG AAAGTCATAT TTCTTCAATT GGTTCGCA CTCAGGGCAA | 3180 |
| GATGGGCGT CGTAGTCCAG TTGGCGATG ATTTCCCTGT GTGTATCCTT ATTGATGATG | 3240 |
| TCTAAATCT GGATATTAGG GTCTTTAATA TCGAGCAGTT TTGTGATAAA ATGTAATTGT | 3300 |
| TCCATATGAA TCTTCTAAT GAGTTGTTTT GTCGCTTTC ATTATAGGTC ATATGGGACT | 3360 |
| TTTTTCTAC AACAAATAG GCTCCATAAT ATCTATAAGG GATTACCCTA CTACAAATAT | 3420 |
| TATAGAGCCG AAAATTCACA TCTAATATAT GCAGACTACT TTGAAATGAA ATTAAAAAA | 3480 |
| TTATTAAAGG ATGACACAAA AGTTTTTGAA AAATCTACAT TCAAATTTGT AGAAGGATAT | 3540 |
| AAAATATACC TGACAGAATC TAAAGAATCT GGAATTAAC AAATGGACAA TGTCATAAAA | 3600 |
| TATTTGAGT TTATTGAATC TAAAAGTATT GCTTTATATT TTCAAAAACG ATTAAATGAG | 3660 |
| CTGATAGATT AAATAGCATT TTCTCTGTG AGATATTGTT TTTAAATAT TGTACTAAAT | 3720 |
| GATTGATGCT ATGTGGAAAT ACAAAAAAAT GTTTTGATA CGAAGTTGAC CTGTATTTTT | 3780 |
| TATACTAATC ATTTTCGTAT TTTTGTATT AAACGATATA AGTTTGTGT AACTTACAA | 3840 |
| GGAATAAAGA CATTAAAAA TAACAGTATA TCTATTGTT TTATATATTT TACGAATTCT | 3900 |
| GCATAAATCT CTTCTAGTA ATGTGTTGTA ACTCTGCTAT AATAGATTTA TTCCTTTTTC | 3960 |
| TGTTTACACA ATTTATTTTA TAGTACCAA AAAGGTCAGG ATTTTGTTCC TGACCTTTGA | 4020 |
| CAACTTTACC GATTCTTTAG TTCTACATAG CGCTTGACC AAATGTTTAC ATAGGCTTCT | 4080 |
| GAGAAAGGAC CACGTCCATT GTTAATCCAA TCAACAAGAA TTTTGACATG TTCTTTTAAA | 4140 |
| ATATAGTCCA AGTCATCAGA ATAATTCATT TTGCGTTTGT GACGCTCGTA CTCTTCAACG | 4200 |
| TCCAAGAGAC GTTTTCCCC ATCTGTAATA ATTTAACAT CCAAATCGTA ATCAATATAC | 4260 |

631

TTCAGTGCTT CTTTCATCCAG ATAGTAGGGG CTAGCCATAT TGCAATAGTA AGAAGTTCCA 4320
 TTATCACGAA TCATGGCAAT GATATTAAAC CAATATTTCT TGTGAAAGTA AACAATAGCC 4380
 GGTTCCTCGAG TGACCCAACG ACGACCATCA CTTTCGGTAA CAAGTGATG ATCGTTGACA 4440
 CCAATAATGG CGTTTCTGT TGTTTTGTAGT ACCATGGTGT CCCGCCAAGT TCGGTGGAGA 4500
 CTCCCATCAT GCTTATAACT TTGAATTGTA ATAAAGTCGC CTTCTTTTGG AAGCTTCATA 4560
 ACTAACCAAC TTTCTACAAT TTATAAGTTT ATCATTACT ATTGTACCAT AAAATTACCC 4620
 AAAATCTGTG AATTTCACTT GGAAATATTA AAGATATTCT CTAAGAGCGC TTGCTATATC 4680
 CGAAAAATCG TAGCCCTTTC GTGCTAAAAC TTGAGTTAAA CGCTGCTTCA GTTCGTATCC 4740
 TTCATACTTT CGGGCATACT TAGTATATTG CTTATCAAGT TCCTTGAAGA TGAGTTCCTG 4800
 AGTCGTTTCT TCATCAACTT GACTATCCAA TTCGTCAAAG GCAATTTTAG CATCAAATA 4860
 AGAGAAGCCC TTGTTAGTCA AGTTCTGGAT AATCTTATCT TGCAGGGCAC GAGCTGGAAG 4920
 TTTTCCCTCA TATTTTTTCA ATAGTTTATT GGCTACACGT TGAGCAACTT CCGAAAAATC 4980
 AAAATCATTC AAGATTCTT CTATAGTAGA TTTTGAAATT CCTTTTGTG CTAATTTCTG 5040
 AGTCAGTACA TAAGGTCCCT TGTCTCCTGA AAGTTGATTG GCATTGATGA TAGCATAAGC 5100
 GTACTGGCTA TCATTAATCC ACTTCTCTTC TTAAAGATTA GCAATGACTT GAGAAACGAT 5160
 GTTTTCATTA ATATCATATT TTTTCAGATA TTCTCTGACC TCTTTTTCAG TACGTGCTTT 5220
 AAAGGATAAG TGGTAGAGGG CCAGATTCTT ACCATAAGAA AATTGAGCAA ACTCTTGAAT 5280
 CTCTTTCAAT TCCTCTTCGC TTATCACCTT ATCTCTCGAT AACATAAAAC GAACAATTGT 5340
 GTCTTCGGTG ATATAGCATT TGTCG 5365

(2) INFORMATION FOR SEQ ID NO: 78:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3636 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 78:

TTTCCAGAAA GAAGTTGAGT AAAGTCTTTA TCAAAGAGAA TGACTTCCGT ATTGGAAGTG . 60
 ACATTAGGTT TTATTCTAC TTTACTAGCG TCCGCCCTAG CATTTTCTAA ATCTTTAATC 120
 TCTTCTGTG CCCTATTTAT AGCCAGCTGA ATAACTGCTT GAGGATTTTC ACTCAGTCCA 180
 TGAAGCTTAT CGTCCACCGA AGTATAAAGA CTCGAATGCA TGACTTGTA AATAATCAGA 240

632

| | | | | | | |
|-------------|------------|------------|------------|------------|------------|------|
| GTCATTGTAG | AAAAAATCAG | GGTGAAGACA | CCGAAGTTGC | GGATAAAATA | ACTAAAGTCA | 300 |
| TCCGCATACC | ATGTTTTTTT | AAGTTTACTG | AACATCTTTT | AAAAGATACC | CAACACTACG | 360 |
| CAAAGTTTGC | AAATTCTCTG | CAAAAGTGGT | TCCCTTTAAT | TTCTTACGGA | CTTTTGAAAC | 420 |
| ATAGACTTCG | ACAACCGAAA | TCGTTGTATC | ACTATCAAAT | CCCCATAGAC | GGTCAAAAAT | 480 |
| CTGCGTCTTA | GGCAAAATCA | CATTTTGATT | TTGAAGGAAA | TAAACTAGTA | AATCGAACTC | 540 |
| TTTCCCCAGC | AATTCGACAG | GAGTATCTTC | AACTTTAACG | GTATTGGTTG | ATAAATTAAC | 600 |
| CACGATATTC | CCATAAGTCA | AGGTGPTTTC | ATTAAACTTC | CCTGAACGTT | TGAGAAGGGC | 660 |
| CTGAATCCGC | ATTTTAAGTT | CTTCTAGGTA | GAAAGGTTTG | GTCAGATAAT | CATCCGCTCC | 720 |
| CAGTTCAAAT | CCATGTCCCT | TGTCATCCAA | ACTTTCCTTG | GCAGTCATAA | TCAGAACTGG | 780 |
| TGTCGTAATT | CCCTTTTCAC | GCAATTCTTT | TAAGACTTGG | AAACCATTTT | TTTCTGGCAA | 840 |
| CATCAAATCC | AGCAAAATCA | AGTCATAGAC | ACCACTCTCA | GCTTCGTAGA | GACCTTCTTC | 900 |
| TCCATCAAAT | ACCTGCATAA | CATCCGCAAA | ATCGTCTAAA | AAGTCAAATA | CTGAATTTGA | 960 |
| CAGACCTAGG | TCATCCTCAA | CCAATAAGAT | TTTATCATG | AGAACTCCT | CCTTATTAAA | 1020 |
| ACTATTATAC | CAAATTTGCC | TTAAAAAAA | CTCAACTCTC | TGCATTTTAC | ATGAGATAGC | 1080 |
| TGAGTTTCT | TTTTATTTTA | GGCTTATTTA | TGCATTTCCG | TATTGAAGAA | CAACTGCTTC | 1140 |
| GACTGCAGCT | TTTTCACGGC | TAATCAAGTC | AACACGCGCT | GCAATTTCCT | TGATTCCCAT | 1200 |
| ACCGATGTTA | CGGCTAAGAG | CAAGGTCAGA | AAGTTGCGGT | TCAAAGAACT | CCTTGATATC | 1260 |
| CGCCAAGCGT | TGCTGAGTCT | TAAATACATG | AGCAGGAAGG | ATAACAAAGC | TATCAAAGCT | 1320 |
| CATATCTCCT | CCAAGGGCTG | CCTTAATCCA | AGCCAGTTT | TCACGCGCCC | AAGACCAAGC | 1380 |
| TGTTTTCTGA | GTTCCTTGAT | GAGCTAGGAA | TTGGTAATAC | CAAGCAGACA | AGTCCTGTGG | 1440 |
| TTTGACCACA | AATTGTCTCT | TCCAAGAAGT | AATCAGGTTT | TGGATATTAT | CCGCATCTGT | 1500 |
| ACTGTATGCA | AGAGCTGCTG | CCAACGGCG | TTTAAAGACA | GCATCTGTTG | CGTGAGTATA | 1560 |
| AGTATCAAGA | TAAAGTGCTA | ACAAGTCTTT | AGTCTCATGA | TGTTTCATCT | CATTAATCAG | 1620 |
| AAC TTGTGAG | CGAATAGCTG | CTGGGAGTCC | TGCAAGATTC | TCCTTGTTGT | TTGCGAAGAT | 1680 |
| TTGGCTAGCG | ACTTGACTAG | CTTCTGCATC | ATTTGAGCGA | ATCATCATCG | AAACAGCCAG | 1740 |
| CTGACGAACC | AATTCATCCT | CATCTGATTC | TCCGTCTTTA | GCTTCAAAAC | CAAGACGGTC | 1800 |
| ATAGTTATGA | CGAGCCAATT | TAGCAACCAG | TCCTTTGAAG | GCTGTTTCAG | CATCCGTTCC | 1860 |
| TTTATCAATA | AAGCGCTCAA | GGGCTGAAAT | CACTTGAGAA | ACAGCTGAAA | CCACCAGATA | 1920 |
| AGACTCTTCC | TTAGCAAGTT | TATCAAGAAC | TGGAAGCAAG | TCTGCATAAG | AAATGTGCCC | 1980 |
| TGCCTCAGCC | AACAAACGAC | GTTCTTGAAC | AATTTGCAGT | TTGCTTGTGT | TATCAAGTGT | 2040 |

633

| | |
|--|------|
| CTCTAGCTCA GCAAGAACAG CTGCTAACAA GTCTCCTTGA TAGTCGGTAA TATAGTGGGC | 2100 |
| AGTATTTTCA GTGTTGAGAC GAAGAGCTCC TTCATTTTCA GCAAGAAGAG CTGCGTAGCC | 2160 |
| AGGGATTTCG ATACTTTCAG TTTCGAGTGT ATCAGGCAAG CCTTTCAGT TGCTATTGAG | 2220 |
| GGGCACCACC CAGAGACGGT TCTTGCTTTC GTTCTCACCG ATGAAGAATT GTTTTGTGA | 2280 |
| AATCTTCAAG ACATCATTTT CAACTTTAAC AGTAAGAACT GGGTAACCAG GCTGTTCCAA | 2340 |
| CCAAGAATCC ATGAAGGCTG CGACATCACG TCCTGACGCT TGACCAAGGG CATCCCAAAG | 2400 |
| GTCACTACCA ATGGTGTTGC TGTATTGGTG TTTTTCAAAG TAGGCGTGCA AACCTTTAGC | 2460 |
| AAAATCAGCA TCTCCTAGCC AACGGCGAAG CATGTGCATG AGACGGCTTC CTTTGGCATA | 2520 |
| GACGATAGCG CCGTCAAAGA GTGTATTGAT TTCATCTGGA TGTTTAACTT CGACGTGGAC | 2580 |
| AGACTGAACG CCATCAGTAG CGTCACGTTT AAGAGCAAGA GGTACTCCAC CTGTTTGGAA | 2640 |
| ATCTTCAAAG ATATTCCAGC TTGGTTCGAT GGTATCCACA CAGACGTATT CCATCATATT | 2700 |
| AGCGAAACTT TCATTGAGCC AAAGGTCATC CCACCATTTT ATAGTCACGA GGTTCCTAAA | 2760 |
| CCATTGGTGA GCCAATTCAT GGGCCACAAC AAGGGCAACT TGTGACGGC TAGCAAATGT | 2820 |
| AGAGTTCTCA TCGACAACCA AGTAAACTTC ACGGTAGGTC ACAAGACCCC AGTTTTCAT | 2880 |
| AGCACCAGCT GAGAAGTCAG GAAGGGCGAT GTGGAGAGAT TGAGGAATTG GGTACTTAAC | 2940 |
| TCCATAGTAA TCTTCGTAAA ACTCGATAGA GCGAACAGCG ATATCCAGTG AGAAATCAAG | 3000 |
| ATTTGAAAGT GGATGTGCTT TGGTTGAGTA GACACCTACC AGGTACCAT TTTTAGTTTT | 3060 |
| AGCGGTCAAC CCTTGCAAAT CACCAGCAAC AAAGGCCAAC AAGTAAGAAG ACATGCGAGG | 3120 |
| TGTTGTCTCA AACTTCCAGA TACCTGTTTC CTTACGGTTT TCAACATCGA TTTCTGGCAT | 3180 |
| GTTTGACAAG GCCAATTCAC CTTCTGCTTG GTCAAAGCGA AGAGAGAGGT CAAAAGTTGC | 3240 |
| TTTGGCTTCA GGCTCATCCA CACATGGGAA AGCTTCGCGC GCAAAATGGC TCTCGAACTG | 3300 |
| AGTAGACAAG ACCTCCTTCT TGACTCCATC AACTGTATAA TAAGAAGGGT AAATCCCTGT | 3360 |
| CATGTTGTCT GTAATTTTAC CAGAAAAGGC AAGAACCAAT TCAACTTGAC CAGCCTCAGC | 3420 |
| CAATTGCGATA TGAAGGGCTT CATTGTCATG GTCAACTGTA AATGGACGAG CTTGACCTGC | 3480 |
| AACTTCTACA GAGGTGATTT CCAAATCTTT TTGGTGGAGG GAGATGCGGT CACTCTGTGC | 3540 |
| TTGACCAGTG ATGGTCACTT TCCAGAAAA AGTCTTGGTC TCACGACTCA AATCTAAAAA | 3600 |
| TAAATCATAA TGTTCAAGAA CAAATTGCTT AATGGG | 3636 |

(2) INFORMATION FOR SEQ ID NO: 79:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5066 base pairs

634

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 79:

| | |
|--|------|
| ATAGCGTGA ATAATCGATT TTAGAGGTAC CATAAGCCAC CTCCTACAAA TAGAAACCGA | 60 |
| TATAAATCAA TGCCTTCCAC CCTTAGACTT CCCTAGTTCC TGTCTCAAGC GAAACATTTT | 120 |
| TTTGAAACAG GAATAAGTTA ACCAATTCAT ACCAATAGCT AGCAGAATAA AAAGAAACCA | 180 |
| AATGCCCCAT AACTTGATAT CTGTCACATT TCTCAAGACG GTATTGAAAA ACAGAACTGA | 240 |
| AACAACGTG CAAAGCAAGGC TAAAAAGAGA ATAGAAGGGG ATGTAAAACC AGTAAAAATA | 300 |
| ATAAAAAATT GGAAAAAACT TACTATTCTT GTTGGCCTTT TCAATCCAGT TATCAAAATA | 360 |
| AAAGTACGGT GCTAAAAGTA AGAATTTAAA CAAATGTTCC ATCACCAGCA TCCCCCTTC | 420 |
| TTTTGATAGC GTTTTCTATT ATTTTATTAT ATCAAAAAA TCCGGAACGT TCATTCCAGA | 480 |
| TTCTACTTTT TTATTGCGT TTTCTTGCGA TGAGATGAAT CGGTGTTCCC TCAAAAACAA | 540 |
| AGGCCTTGCG GATTTGATTT TCCAAGAAAC GCAGGTAAGA AAAGTGATG AGTTCTTCTT | 600 |
| CATTGACAAA GATGACAAAG GTTGGTGGTT TGGTTGCCAC TTGGGTCGCA TAGAAAATCT | 660 |
| TGAGACGTTT TCCTTTGTCT GTCGGTGTG GGTGATGGC AATGGCATCC ATGATGACAT | 720 |
| CGTTCAAGAC AGCTGATGGA ATACGTGTAT TTTGACTTTC GCTGATTTGC TTAATCATCT | 780 |
| CAGGAAGTTT GTGGAGACGT TGCTTGGTTA AAGCTGATC AAAGATAATC GGTGCGTAAG | 840 |
| GCAGGTATTG GAACTGCTCA CGGATATCTT CTTCCTCAGT TTTCATAGTG TGGTTATCTT | 900 |
| TTTCAAGCGT ATCCCCTTG TTGACCACGA TAATCATCCC TTTACCAGCT TCATGGGCAA | 960 |
| ATCCTGCGAT ACGCTTGTCG TACTCACGAA TGCCTTCTTC CGCATTGATG ACCATCAAGA | 1020 |
| CCACATCTGA ACGGTCAATA GCACGCATGG CACGCATAAC AGAGTATTTT TCAGTATTTT | 1080 |
| CATAAACCTT ACCAGACTTA CGCATACCAG CCGTATCAAT CATGGTAAAC TCTTGACCAT | 1140 |
| CTGTATCTGT AAAGTGGGTA TCAATGGCAT CACGAGTTGT TCCAGCAACA GGACTAGCAA | 1200 |
| TAACACGGTC TTCTCCCAAG ATAGCATTGA TCAAGCTTGA TTTTCCAACG TTAGGACGAC | 1260 |
| CAATCAAGCT AAACCTTAATG ACATCTGGAT TTTCTTCCTC ATATTCATTT GGAAGATTTT | 1320 |
| CTACGATCGC ATCTAGCACA TCCCCTGTAC CGATTCCATG GACAGATGAG ATAGGCAATG | 1380 |
| GTTCACCCAA ACCGAGAGCA TAGAAATCAT ATATATCATT TCTCATCTCA GGGTTGTCCA | 1440 |
| CCTGTGTTGAC TGCGAGGATA ACTGGTTTGT GGTCTTATA AAGCTTACGA GCTACGTATT | 1500 |
| CGTCTGCATC AGTAATTCCT TCCTTACCAG ACACGACAAA AACGATAACA TCTGCTTCTT | 1560 |

635

| | |
|--|------|
| CCATGGCAAT TTCTGCCTGG TGCTTGATT GTTCCATGAA AGGAGCATCG ACATCATCAA | 1620 |
| TTCTCTCCTGT ATCAATCATG CTAAAAGAAC GATTGAGCCA CTCACCCGTT GCATAAATAC | 1680 |
| GGTCACGTGT CACTCCTTCG ACATCTTCTA CAATGGAGAT TCGCTCACCA GCGATCCGAT | 1740 |
| TAAATAGGGT TGATTTCCCA ACATTGGGAC GTCCTACAAT GGCAATAGTT GGTAGGGCCA | 1800 |
| TAATTTCTCA CTTTCTACAA TAATTTCTTC TGTTCAGAT TTTTCTAGT TGAGCTTGGT | 1860 |
| TCAGCTTGAC CAAACTGTTT TGCTAGGCGC TGACTCCAGC TTGTGGTCGC ACGCGCCCCA | 1920 |
| GCATAGTCAG CCTGAACACG GTCATAAGCT TGGATTGCCT CAGTTGACTG TTCTTGGTAT | 1980 |
| TCTTCTCAA AGACAACATT CTCTAGTGGC AGTCTCGGTT TCATATCATG ATGTTGATTT | 2040 |
| GGCACACCCA GTGCCATCCC AAAGACAGAA TAGGTGTAGT CAGGTAGGTT AAAGAGCTCT | 2100 |
| GCCACTTCTT CAGACTTGTA TCGAACCAAA CCGATAATCA CACCACCATA GCCCAAGCTT | 2160 |
| TCAGCTGCCA ACAAGGCGTT TTGTCCAGCA AGAGCTGCAT CGACCGAACT AATCAAGAGA | 2220 |
| CCTTCCACAC CTTGGGGTTG GAAGGTGTCG GTATGAAGTC GGGCTCCCTT TTCTGCTCGG | 2280 |
| TTCAAATCTC CGACAAAGAG AAGGAAAACA GCAGACTGGC GAATGGCTTC TTGAGGTACC | 2340 |
| AATTCATACA AGGCATCTTT CTTCTCTTGA CTTCTGACCA CAATCACAGA GTAGGATTGG | 2400 |
| AAATCTTCC AAGATGATGC CATCTGGGCT GCTGTCAAAA TCTCATTTAA GTCTACTTGG | 2460 |
| GGAAATTTCTT GCTCTTTAAA CCTGCGCACT GAAGTATGAG CCTTCATCAA TTTAATGGTT | 2520 |
| TCTGTATCG ACGGTTTACT CCTTCTAAAC GAGTCTCCTC AGCCAAATAA CGGATGCGTT | 2580 |
| CCATGACCCG TCTGGCTTCC CAGGTTTCGT CATTTCCATG TTTCACTTTC GCAAAATGCT | 2640 |
| TCTCCAAATC TTCAAAGTTG AAGTTGGATG TGAAAAAGGT CGGTAAATTT TCCTGCATCC | 2700 |
| GATATTGGAG AATGACCTGC AGGATTTTCGT CACGCACCCA AACGGTTGAT TGCTCGGCGC | 2760 |
| CAATATCATC TAAAATCAGG ACCTCAGACA GCTTAATCTC ATCCACCAAG GTCTTAACAT | 2820 |
| TGCCATCACT GATAGCATTT TTGACATCAA TGACAAAGCT AGGATAGTGG AGGAGAGTTG | 2880 |
| ATGAAACACC ACGTTTTTCT GATAAATCAT GAGCTAAGGC CGCCACCATG AAATTTTAC | 2940 |
| CCACACCAAA GTCTCCATAT AAGTAAAGAC CTTTTCGAAT AGCTGGATAT TGCTCCACGA | 3000 |
| AGGCTAGTAG CTTTTCAAAA ACTGGTAAGC GCCCCAAATC ATCCAAGTCA ACTTGAGCCA | 3060 |
| AACTAGCTTT CTTGAGACTG GCTGGTAGAT TGATTAACTT GAGACGGTTC TTAATAGCCG | 3120 |
| CTTCTTTTTC AGCCGCGATT AGCTCAGGAG TTTCTTCATA TGAAACATCT GCATAACCAT | 3180 |
| GATTCTTAAC CAAAATCGGC TTGTAGCCTT TGGCAATATA ATCCGTATCC CCACGGAGAA | 3240 |
| ACTGTGCACG CTCGGTGATG TACTGATTAA ACTTGGAGAT ACTGCGATTT AATTCCTTTG | 3300 |

636

| | |
|--|------|
| GAGTTAAGGA TTCTTGCTGG ATAAAGGCCG CAACATCAGG GTCCTTCATG ATTTTCTGGA | 3360 |
| CCAAATCTTG ATAATAAAAA CGGCTGGGTT GACGTTTGAG TACGTCTCCG AACTTTTCCA | 3420 |
| TCTAATCTCC TCCTTTTCT AATCGAGCTA ATAGTTCTTG CTCTTACGT TCTAGTTCCA | 3480 |
| GACGAGTTTC CTCGCTGGTT TCATTCTTAT ATTGAGGATT ACTCCATTTA GGAACATTGG | 3540 |
| TTTTTTCTGG GGCAGTCTGA TTCTGTTTTT GTGTTTTTGC TTTCTGCCCT CGATCACGAA | 3600 |
| TTCTGTA AAC GGCCTCTTCT GCCGAATGAA TCTTTTGATA GGCATAGTCA TTGGCTACCT | 3660 |
| TCATGGCATA TTTCTCATTG ATATTTGCCG AATCCACCTT ATTAAGGTC AATAAGAGAA | 3720 |
| TAATATTGAT GACTTCGTCC AGTAAGCCCA AGCCAGCCAT CTGTTGCAAG AGTTCTCTTT | 3780 |
| CTGTTTGGGT AATGGTTCCC TTGCGTGTTC GCTTGATTTC TGCTAAGAAC TGCAGGGCAG | 3840 |
| TTTTACTTTT AGCTTCTTTG ATAATGGTCG CTTCTTAAG ACTAAAGTCA GAGGAACTG | 3900 |
| GTTTTTGAGC AATTTTTTCA CGCATGCGTT TGGTTGAAAT AACCTGGGAA ACAGCTGTTC | 3960 |
| ACTTGCCCAA TTGATAGGTT TCAAACCAAG TCCATTTCTT CTCCTCGGCA ATAGCAAAGA | 4020 |
| GGTTTAAGAC ATCGGACTGC TCATCCGCAA AACGAAGTCC ATCTCGAGCC ATCAGCTGGC | 4080 |
| GAAAATGTTT CAAGTCAAAA TCATTGGCCA CTTTCTTCTT GAGACCAAGG TCTTCTTGAC | 4140 |
| TGCCTAGTTC TGCCAATCTT GGAAAGACTT GATTGAGTGA GACAGGTATT TCTTCACCAT | 4200 |
| CAGCACTTTC AACTTTCAAA TCCTCCACAG CTACATCGCC AATCTTTTTC TCTAAGAGTC | 4260 |
| TGCGATAAAC AGGATGCCCC AAGAAGTCTT GACTAGATAG AGGAGCATGG AGGGCTAGCT | 4320 |
| GATAAACATC ACCCTTTTGA TAGAGGTCA AGAGATTAAA AGCAGATAAG ATTTTCAATG | 4380 |
| ATTTTATCAG TCTATCCATC CCAAAGTTGA GATGGTTGAG AATGCTTGAA AAAAGATATT | 4440 |
| CCTTTCTACC ATTATCCCAA AAAGTATTG TATAAGATA AAGGCTCAGT GCCTCCTGAC | 4500 |
| CGATAATCGG GAGGTAGCAC TGTACCAGAG ATGAGGTATC TTGCGACACC CGATTATTCT | 4560 |
| TTAGATAAGA AAAACGGTCA ATTGGCTTCA TTTATCTTTC CTTTCTCTT TTAGAGGACT | 4620 |
| GGGTGATTTG TTGGAGCAAG CTCTCTAACT CACTGACATC CTTAAACTA CGATAGACAC | 4680 |
| TAGCAAAACG TACATAGGTA ATCTCGTCCA ATTCAGCCAA CTCCTCCATG ACGAGTGAAC | 4740 |
| CAATGTCCTC ACTTTGAATT TCATTTTCAT TTCGACCACG GAGTTTCTGT TCGATACGAT | 4800 |
| TGACTACCAT GTTGATTTC TCACTTGACA CAGGACGTTT CTGGGCTGAG CGGATAATCC | 4860 |
| CATTAAAGAT TTTATCTCTG GAGAATTGTT CCCGTGTGCC ATCTTTTTTA ACAACCACTA | 4920 |
| AGGTTCTTTC TTCTACTCGT TCGTAGGTTG TAAAACGGTG TTGGCATTTC TCGCACTCAC | 4980 |
| GTCTTCTACG AATGGTGTTC CCTTCTTCTG CTTGGCGACT ATCGATAACA CTTGACTTGG | 5040 |
| TAGCCCCACA TTTTGGACAG GGTACC | 5066 |

637

(2) INFORMATION FOR SEQ ID NO: 80:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 9607 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 80:

| | |
|--|------|
| CACCTGAAGT ATTTGAAACA GCTATGGAAA ACATCATGCC TGTACTTGAA GTACGTGCAC | 60 |
| GTCGTGTTGG TGGTCTAAC TACCAAGTCC CAGTTGAAGT TCGTCCAGAA CGTCGTACAA | 120 |
| CACCTGGACT TCGTTGGTTG GTAACAATCG CTCGTCTTCG TGGTGAACAC ACAATGCAAG | 180 |
| ACCGTCTTGC AAAAGAAATC TTGGATGCTG CTAACAACAC TGGTGCAGCA GTTAAGAAAC | 240 |
| GTGAAGATAC TCACCGTATG GCTGAAGCTA ACCCTGCATT CGCACACTTC CGTTGGTAAG | 300 |
| ATAGGATGCG AAAGCGTTAA GAAAGTCCCA GAGAAAATAG GGAATCGAAG CAGGTTGCGG | 360 |
| TTGCAACCAA TGAGATTCAT CTTTTCTCC AGACTTTTAG CTTGAGCTCA ACTAAATCAT | 420 |
| GATGCTAGGA ACGGTAAGGA TGCAAGGTAA AAATAGGAAA CTGACGCAGT ATTCGACGAA | 480 |
| TACAAGGAGT TTTATCTTTT TCACGCAGCA TCCCGTTCCA GCTCACATCG GCTAACTAAC | 540 |
| TTTAGCCCGG GTTCAAATTA GCTAAATCGA TTAGTATTAG CTATAACTCA GCTTACCATC | 600 |
| TCGTAAGTTG AAACCAACAA TAGCATGAAA ACATTGAGAA CGGGTAGGTC CTGCCTATCC | 660 |
| GTTTTTATTA AAATCGTGTT ATAATAGAAT AGAAATCAAA AATAAATAGG AGAAACAAAC | 720 |
| CTCATGGCAC GCGAATTTTC ACTTGAAAAA ACTCGTAATA TCGGTATCAT GGCTCACGTC | 780 |
| GATGCCGGTA AAACAACAAC TACTGAGCGT ATTCTTTACT ACACTGGTAA AATCCACAAA | 840 |
| ATCGGTGAAA CTCACGAAGG TCGGTCACAA ATGGACTGGA TGGAGCAAGA GCAAGAACGT | 900 |
| GGTATCACGA TCACATCTGC TGCGACGACA GCTCAATGGA ACAACCACCG CGTAAACATC | 960 |
| ATCGACACAC CAGGACACGT GGACTTCACA ATCGAAGTAC AACGTTCTCT TCGTGTATTG | 1020 |
| GATGGTGCGG TTACCGTTCT TGA CTCACAA TCAGGTGTTG AGCCTCAAAC TGAAACAGTT | 1080 |
| TGGCGTCAAG CAACTGAGTA CGGAGTTCCA CGTATCGTAT TTGCCAACAA AATGGACAAA | 1140 |
| ATCGGTGCTG ACTTCCTTTA CTCTGTAAGC ACACTTCACG ATCGTCTTCA AGCAAATGCA | 1200 |
| CACCCAATCC AATTGCCAAT CGGTTCTGAA GATGACTTCC GTGGTATCAT TGA CTTGATC | 1260 |
| AAGATGAAAG CTGAAATCTA TACTAACGAC CTTGGTACGG ATATCCTTGA AGAAGACATC | 1320 |
| CCAGCTGAAT ACCTTGACCA AGCTCAAGAA TACCGTGAAA AATTGATTGA AGCAGTTGCT | 1380 |

638

| | |
|--|------|
| GAAACTGACG AAGAATTGAT GATGAAATAC CTCGAAGGTG AAGAAATCAC TAACGAAGAA | 1440 |
| TTGAAAGCTG GTATCCGTAA AGCGACTATC AACGTTGAAT TCTTCCCAGT ATTGTGTGGT | 1500 |
| TCAGCCTTCA AAAACAAAGG TGTTC AATTG ATGCTTGATG CGGTTATCGA CTACCTTCCA | 1560 |
| AGCCCACTTG ACATCCCAGC AATCAAAGGT ATTAACCCAG ATACAGACGC TGAAGAAATT | 1620 |
| CGTCCAGCAT CTGACGAAGA GCCATTTGCA GCTCTGCCT TCAAGATCAT GACTGACCCA | 1680 |
| TTCGTAGGTC GTTTGACATT CTCCGTGTT TACTCAGGTG TTCTTCAATC AGGTTTATAC | 1740 |
| GTATTGAATA CTTCTAAAGG TAAACGTGAA CGTATCGGAC GTATCCTTCA AATGCACGCT | 1800 |
| AACAGCCGTC AAGAAATCGA CACTGTTTAC TCAGGTGATA TCGCTGCTGC CGTTGGTTTG | 1860 |
| AAAGATACTA CAACTGGTGA CTCATTGACA GATGAAAAAG CTAAAATCAT CCTTGAGTCA | 1920 |
| ATCAACGTTT CAGAACCAGT TATCCAATTG ATGGTTGAGC CAAAATCTAA AGCTGACCAA | 1980 |
| GACAAGATGG GTATCGCCCT TCAAAAATTG GCTGAAGAAG ATCCAACATT CCGCGTTGAA | 2040 |
| ACAAACGTTG AAATGGTGA AACAGTTATC TCAGGTATGG GTGAACCTCA CCTTGACGTC | 2100 |
| CTTGTGATC GTATGCGTCG TGAGTTCAAA GTTGAAGCGA ACGTAGGTGC TCCTCAAGTA | 2160 |
| TCTTACCGTG AAACATTCCG CGCTTCTACT CAAGCACGTG GATTCTTCAA ACGTCAGTCT | 2220 |
| GGTGGTAAAG GTCAATTCCG TGATGTATGG ATTGAATTTA CTCCAAACGA AGAAGGTAAA | 2280 |
| GGATTGCAAT TCGAAAACGC AATCGTCGGT GGTGTGGTTC CTCGTGAATT TATCCCAGCG | 2340 |
| GTTGAAAAAG GTTTGGTAGA ATCTATGGCT AACGGTGTTC TTGCAGGTTA CCCAATGGTT | 2400 |
| GACGTAAAG CTAAGCTTTA TGATGGTTCA TATCACGATG TCGACTCATC TGAAGTACC | 2460 |
| TTCAAGATTG CGGCTTCACT TTCCCTTAAA GAAGCTGCTA AATCAGCACA ACCAGCTATC | 2520 |
| CTTGAACCAA TGATGCTTGT AACAACTACT GTTCCAGAAG AAAACCTTGG TGATGTTATG | 2580 |
| GGTCACGTAA CTGCTCGTCG TGGACGTGTA GATGGTATGG AAGCACACGG TAACAGCCAA | 2640 |
| ATCGTTCGTG CTTACGTTCC ACTTGCTGAA ATGTTCGGTT ACGCAACAGT TCTTCGTCT | 2700 |
| GCATCTCAAG GACGTGGTAC ATTCAATGATG GTATTTGACC ACTACGAAGA TGTACCTAAG | 2760 |
| TCAGTACAAG AAGAAATTAT TAAGAAAAAT AAAGGTGAAG ACTAATCCGT CCTCACTCTA | 2820 |
| GAAGGAAGTC ACTTAGTGGC TTCCTTTTGT CTTTAGAAAA TACCTCTAAA TATGGTAAAA | 2880 |
| TAGTAGAAGA ATAATGTGAG GAAAATGAAT GTCAAATAGT TTTGAAATTT TGATGAATCA | 2940 |
| ATTGGGGATG CCTGCTGAAA TGAGACAGGC TCCTGCTTTA GCACAGGCCA ATATTGAGCG | 3000 |
| AGTTGTGGTT CATAAAATTA GTAAGGTATG GGAGTTTCAT TTCGTATTTT CTAATATTTT | 3060 |
| ACCGATTGAA ATCTTTTATAG AATTAAAGAA AGGTTTGAGC GAAGAATTTT CTAAGACAGG | 3120 |
| CAATAAAGCT GTTTTGTAAA TTAAGGCTCG GTCTCAAGAA TTTTCAAATC AGCTCTTGCA | 3180 |

639

| | |
|---|------|
| GTCCTACTAT AGGGAGGCTT TCTCTGAAGG TCCATGTGCT AGTCAAGGTT TTAAGTCCCT | 3240 |
| TTATCAAAAT TTGCAAGTTC GTGCTGAGGG TAATCAGCTA TTTATTGAAG GATCTGAAGC | 3300 |
| GATTGATAAG GAACATTTTA AGAAGAATCA TCTTCCTAAT TTAGCCAAAC AACTTGAAAA | 3360 |
| GTTTGGTTTT CCAACTTTTA ACTGTCAAGT CGAGAAGAAT GATGTCCTGA CCCAAGAGCA | 3420 |
| GGAAGAGGCC TTTCATGCTG AAAATGAGCA GATTGTTCAA GCTGCCAATG AGGAAGCGCT | 3480 |
| CCGTGCTATG GAACAACTGG AGCAGATGGC ACCTCCTCCA GCGGAAGAGA AACCAGCCTT | 3540 |
| TGATTTTCAA GCGAAAAAAG CTGCAGCTAA ACCCAAGCTG GATAAGGCGG AGATTACTCC | 3600 |
| TATGATCGAA GTGACGACAG AGGAAAATCG TCTGGTATTT GAAGGGGTTG TTTTGTATGT | 3660 |
| GGAGCAAAAA GTGACTAGAA CAGGTCGTGT TTTAATCAAC TTAAAAATGA CGGACTATAC | 3720 |
| TTCAAGTTTT TCTATGCAAA AGTGGGTAA AAACGAGGAA GAGGCCCAGA AGTTTGACCT | 3780 |
| CATCAAGAAG AATTCTTGGC TCCGAGTTCG AGGGAATGTG GAGATGAATA ACTTCACACC | 3840 |
| CGATTTGACT ATGAACGTAC AGGATCTGCA GGAAGTTGTT CACTATGAGC GGAAGGATTT | 3900 |
| GATGCCAGAA GGTGAGCGTC GGGTTGAGTT TCATGCTCAT ACTAACATGT CGACTATGGA | 3960 |
| TGCTTTGCCA GAGGTCAAG AGATTGTGTC AACAGCTGCT AAGTGGGAC ACAAGGCGGT | 4020 |
| TGCTATCACG GACCATGGGA ATGTCCAGTC CTTTCCACAT GGCTATAAGG CGGCTAAGAA | 4080 |
| AGCGGGAATC CAGCTGATCT ATGGGATGGA AGCCAATATC GTGGAGGACC GTGTCCCTAT | 4140 |
| CGTCTATAAC GAAGTGGAGA TGGACTTGTC AGAAGCAACC TACGTGGTCT TTGACGTGGA | 4200 |
| AACGACGGGA CTTTCAGCTA TCTATAATGA CTTGATTCAG GTTGCGGCTT CTAAGATGTA | 4260 |
| CAAGGGGAAT GTTATTGCTG AATTGTATGA ATTTATCAAT CCTGGGCATC CCTTGTGAGC | 4320 |
| CTTTACTACA GAGTTAACTG GAATTACAGA TGATCATGTC AAAAATGCCA AACCACTAGA | 4380 |
| ACAAGTTTTG CAAGAATTCC AAGAATTTG CAAGGATACG GTCCTAGTTG CCCACAATGC | 4440 |
| TACCTTTGAC GTTGGCTTTA TGAATGCTAA TTATGAGCGG CATGATCTTC CAAAGATTAG | 4500 |
| TCAGCCAGTT ATTGATACGC TGGAGTTTGC TAGAAACCTC TATCCTGAGT ATAAACGCCA | 4560 |
| TGGTTTGGGG CCTTTGACCA AGCGTTTTGG TGTGGCCTTG GAACATCACC ACATGGCCAA | 4620 |
| CTACGATGCG GAAGCGACTG GTCGTCTGCT TTTCATCTTT ATCAAAGAGG TAGCAGAAAA | 4680 |
| ACATGGTGTG ACCGATTTAG CTAGACTCAA CATTGATCTA ATCAGTCCAG ATTCTTACAA | 4740 |
| AAAAGCTCGG ATCAAGCATG CGACCATCTA TGTCAAGAAT CAGGTAGGTC TAAAAAATAT | 4800 |
| CTTTAAGCTG GTTTCCTTGT CTAATACCAA GTATTTTGAA GGAGTGCCAC GGATTCCGAG | 4860 |
| AACGGTTCTA GATGCCCATC GAGAGGGCTT GATTTTAGGT TCAGCCTGTT CAGAGGGTGA | 4920 |

640

| | |
|--|------|
| AGTTTTTGAC GTGGTCGTTT CTCAAGGTGT GGATGCGGCG GTTGAGGTGG CCAAGTATTA | 4980 |
| TGATTTTATC GAGGTCATGC CACCGGCTAT CTATGCACCC TTGATTGCCA AAGAGCAGGT | 5040 |
| CAAGGATATG GAGGAACTCC AGACCATTAT CAAGAGTTTG ATAGAGGTTG GAGACCGCCT | 5100 |
| TGGCAAGCCT GTTCTGGCTA CGGGAAATGT TCACTATATC GAACCGGAAG AAGAGATTTA | 5160 |
| TCGTGAAATT ATCGTCCGTA GTTTGGGACA GGGTGCGATG ATTAATCGAA CTATCGGTCA | 5220 |
| TGGTGAAACAT GCCCAACCAG CACCACTTCC AAAGGCTCAT TTTCGAACGA CTAATGAGAT | 5280 |
| GTTGGATGAA TTTGCCTTTT TGGGAGAGGA ACTGGCTCGT AAAGTGGTTA TTGAAAACAC | 5340 |
| CAATGCCTTG GCAGAAATAT TTGAATCCGT TGAAGTCGTT AAGGGTGACT TGTATACGCC | 5400 |
| TTTCATCGAC AAGGCTGAAG AAACAGTTGC TGAGTTGACC TATAAGAAAAG CTTTTGAGAT | 5460 |
| TTATGGAAAT CCGCTGCCAG ATATTGTTGA TTTGCGGATT GAAAAAGAAT TAACATCCAT | 5520 |
| ACTGGGGAAT GGATTTGCTG TGATTTATCT GGCATCGCAG ATGCTGGTGC AACGTTCTAA | 5580 |
| TGAACGGGGT TATTTGGTTG GTTCTCGTGG GTCTGTCGGA TCTAGTTTCG TTGCGACCAT | 5640 |
| GATTGGGATT ACGGAGGTCA ATCCTCTCTC TCCTCACTAT GTCTGTGGTC AGTGTCACTA | 5700 |
| CAGTGAGTTT ATCACAGATG GTTCGTACGG TTCAGGATTT GATATGCCCC ATAAGGACTG | 5760 |
| TCCAAACTGT GGTCACAAAC TCAGTAAAAA CGGACAGGAT ATTCCGTTTG AGACCTTCCT | 5820 |
| TGGTTTTGAT GGGGATAAGG TTCCTGATAT TGACTTGAAC TTCTCGGGAG AAGATCAGCC | 5880 |
| TAGCGCCAC TTGGATGTGC GTGATATCTT TGGTGAAGAA TATGCCTTCC GTGCGGGAAC | 5940 |
| GGTTGGTACG GTAGCTGCCA AGACTGCCTA TGGATTTGTC AAAGGTTACG AGCGAGATTA | 6000 |
| TGGCAAGTTT TATCGTGATG CAGAAGTAGA ACGCCTCGCT CAAGGAGCGG CGGGTGTCAA | 6060 |
| GCGGACAACA GGCCAACACC CGGGGGGAAT CGTTGTTATT CCGAACTACA TGGATGTCTA | 6120 |
| CGATTTTACG CCTGTCCAGT ATCCAGCAGA TGATGTCACG GCTGAATGGC AGACCACTCA | 6180 |
| CTTTAACTTC CACGATATCG ATGAGAACGT CCTCAAAC TC GATGTACTGG GACATGATGA | 6240 |
| TCCGACTATG ATTCGAAAAC TTCAGGATTT GTCTGGTATT GACCCTAATA AAATTCCTAT | 6300 |
| GGATGACGAA GGCGTGATGG CACTCTTPTC TGGGACTGAT GTGCTAGGGG TAACACCTGA | 6360 |
| ACAAATTGGA ACGCCTACGG GTATGTTGGG GATTCCAGAG TTTGGAACAA ATTTCCGTACG | 6420 |
| TGGAATGGTA GACGAAACCC ATCCGACAAC CTTTGCGGAA TTGCTTCAGC TGTCTGGTCT | 6480 |
| GTCCACGGT ACTGATGTTT GGTGCGGAA TGCTCAGGAT CTGATTAAGC AAGGAATAGC | 6540 |
| GGACCTATCG ACTGTTATCG GTTGTGCGGA CGACATCATG GTTTACCTCA TGCATGCGGG | 6600 |
| TCTGGAACCT AAGATGGCCT TTACCATTAT GGAACGGGTA CGTAAGGGTT TGTGGCTAAA | 6660 |
| GATTTTCAGAA GAGGAGAGAA ATGGCTATAT CGAAGCAATG AAGGCTAATA AGGTGCCAGA | 6720 |

641

| | |
|--|------|
| GTGGTATATC GAATCCTGTG GGAAAATTAA GTACATGTTT CCTAAGGCC ATGCGGCAGC | 6780 |
| CTACGTTATG ATGGCCTTGC GTGTAGCTTA CTTCAAGGTT CACCATCCTA TTTATTACTA | 6840 |
| CTGTGCTTAC TTCTCCATTC GTGCTAAGGC TTTTGATATC AAGACCATGG GTGCGGGCTT | 6900 |
| GGAGGTCATC AAGCGCAGAA TGAAGAAAT CTCTGAAAA CGGAAGAACA ATGAAGCCTC | 6960 |
| TAATGTGGAA ATCGATCTCT ATACAACCTT TGAGATTGTC AATGAGATGT GGAACGAGG | 7020 |
| TTTCAAGTTT GGTAAATTAG ATCTCTACTG TAGTCAGGCG ACAGAGTTCC TCATCGACGG | 7080 |
| GGATACCCCTT ATCCCACCAT TTGTAGCAAT GGATGGTCTG GGAGAGAACG TTGCCAAGCA | 7140 |
| ACTGGTGCGG GCGCGTGAAG AGGGAGAATT CCTCTCTAAA ACAGAACTAC GCAAGCGTGG | 7200 |
| TGGACTCTCA TCAACCTTGG TTGAAAAGAT GGATGAGATG GGTATTCTTG GAAATATGCC | 7260 |
| AGAGGATAAC CAGTTGAGTT TGTTTGATGA GTTGTTTTAA AAAATTGCTT AATAATCTAT | 7320 |
| TAAAAGAGGC TAACGTATAT CCAATAGATT TACATTAGCT TTCTTTTGTG TTTAAATAGT | 7380 |
| CTATGGAAAG AGGGTGAGAG TATGTCAAAG ATGAGTATAA GCATCCGCTT GGATAGTGAG | 7440 |
| GTTAAGGAGC AGGCCCAACA GGTGTTTAGT AATCTGGGAA TGGATATGAC AACAGCTATT | 7500 |
| AATATTTTCC TTCGTCAGGC AATTCAATAT CAGGGATTAC CTTTGTGATG TAGACTAGAC | 7560 |
| GAAAATCGGA AGTTGCTCCA AGCGTTAAGC GATTTAGACC AAAATCGTAA TATGAGCCAG | 7620 |
| TCTTTTGAAT CAGTCTCAGA TTTGATGGAG GACTTACGTG CTTAAGATTC GTTATCATAA | 7680 |
| ACAGTTTAAA AAAGATTTTA AGTTGGCTAT GAAGCGTGGT TTGAAGGCAG AATTATTAGA | 7740 |
| AGAAGTTTGT AATTTTCTGG TTCAAGAAAA AGAACATCCT GCCAGAAATC GTGATCATTC | 7800 |
| ATTGACGGCA TCCAAGCATT TTCAAGGAGT TCGTGAATGC CATACCCAGC CAGATTGGCT | 7860 |
| TTTGTTTAT AAAGTAGACA AGTCGGAATT GATTTTAAAT TTGCTGAGGA CAGGCAGTCA | 7920 |
| CAGTGATTTA TTTTAATCTA TTTTAAGGGG GTTCTCATGA AACTAAGAAT ATTTGCGGAA | 7980 |
| GATAAGCCGG CTAAGAAGGT ATTTGAATAT CAATTAGAAC TTGCTGATCG TACAATCTTT | 8040 |
| CTATCGACAG CACTCTTGTG AGGTGCTATT GCTTTAGCAG GAATCTTTTC TGCTTTGAAA | 8100 |
| GAAAAATAAA AATAGAAAAG AGAAAACAGA ATGGTTTAC CAAATTTTAA AGAAAATCTA | 8160 |
| GAAAAATATG CGAAATTGTT GGTTCGAAC GGAATTAACG TGCAACCTGG TCACACTTTG | 8220 |
| GCTCTCTCTA TTGATGTGGA GCAACGTGAA TTGGCACATC TAATCGTGAA AGAAGCTTAT | 8280 |
| GCCTTGGGTG CGCATGAGGT CATCGTTCAG TGGACAGATG ATGTGATTAA CCGTGAGAAA | 8340 |
| TTCTCCATG CCCCAGTGA GCGTTTGGAC AATGTGCCAG AATACAAGAT TGCTGAGATG | 8400 |
| AACTATCTCT TGGAGAATAA GGCTAGCCGT CTGGAGTTC GTTCATCTGA TCCAGGTGCC | 8460 |

642

| | |
|--|------|
| TTGAACGGAG TGGACGCTGA CAAGCTTTCA GCTTCTGCTA AAGCTATGGG ACTTGCCATG | 8520 |
| AAGCCTATGC GTATCGCAAC TCAATCTAAC AAGGTTAGCT GGACTGTAGC AGCTGCAGCA | 8580 |
| GGACTTGAGT GGGCTAAGAA AGTCTTCCCA AATGCTGCGA GCGACGAAGA AGCAGTTGAT | 8640 |
| TTCTTTTGGG ACCAAATTTT CAAAACCTGC CGTGTCTACG AAGCAGATCC TGTTAAGGCT | 8700 |
| TGGGAGGAAC ATGCAGCCAT TCTCAAGAGC AAGGCCGATA TGCTTAATAA GGAGCAATTT | 8760 |
| TCAGCCCTTC ACTACACAGC GCCAGGAACA GATTTAACAC TTGGTTTGCC AAAGAACCAC | 8820 |
| GTTTGGGAAT CAGCTGGTGC TGTCAATGCA CAGGGCGAAG AATTCTTGCC AAATATGCCA | 8880 |
| ACAGAAGAGG TCTTCACAGC GCCTGACTTC CGTCGTGCAG ATGGTTATGT CACTTCTACA | 8940 |
| AAACCGCTTA GCTACAACGG AAATATCATT GAAGGCATTA AGGTGACCTT TAAGGATGGA | 9000 |
| CAAATCGTAG ATATCACTGC TGAGAAGGGT GATCAGGTTA TGAAAGACCT TGTCTTTGAA | 9060 |
| AATGCGGGTG CGCGTGCCTT GGGTGAATGT GCCTTGGTAC CAGATCCAAG TCCAATTTCT | 9120 |
| CAGTCAGGCA TTACCTTCTT TAACACCCTT TTCGATGAAA ATGCGTCAAA CCACTTGGCT | 9180 |
| ATCGGTGCAG CCTATGCGAC TAGCGTTGTT GATGGAGCGG AGATGAGCGA AGAGGAGCTT | 9240 |
| GAAGCTGCAG GGCTTAACCG TTCAGATGTT CACGTAGACT TTATGATTGG TTCTAACCAG | 9300 |
| ATGGATATCG ATGGTATTCG TGAGGATGGA ACGCGGGTAC CTCTTTTCCG TAATGGGAAT | 9360 |
| TGGGCAAATT AAGGAGATAA TATGTTAGGA AGTATGTTCCG TTGGTCTCCT AGTGGGATTT | 9420 |
| TTAGCAGGTG CTATGACCAA TCGTGGAGAG CGAATGGGAT GTTTTGGAAA AATGTTTCTC | 9480 |
| GGTTGGATCG GAGCCTTCT AGGTCCTTG CTCTTTTGAA CTGGGGGCC AGTTTATCA | 9540 |
| GGAACAGCTA TTATCCAGC GATTTTAGGA GCCATGATTG TTTTAGCTAT TTTTGGAGA | 9600 |
| CGAGGAA | 9607 |

(2) INFORMATION FOR SEQ ID NO: 81:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 14231 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 81:

| | |
|---|-----|
| CTACAAGATA ATTCCAGCTA TAACATCCGC TATAATAGTA AGAGCGAGCT CTATGATAAG | 60 |
| GCTCATTAGT TTCACCTCCT CTCACGAACC CATAGGAACG TAATCGGTAA CCGATGACAA | 120 |
| AAATAGTATA CCACAATACA TTTAGATCAT CAAGGTCCT TAATTCTTGA AATATCAGAT | 180 |
| CTAAGAGAAA AATCTTTAAA ATCAGAAAAA CGCATAATAT CAGGTGTGCA AAAACTTGAT | 240 |

643

| | |
|--|------|
| ACTATGCGTT TTATTGTGGG AAGGTTTACT CCATTTTCTC CTGAAATTGA GTTTTTGTCC | 300 |
| AGCCTCTGTT TTTAGGGTTG CTAAGAAAAT AATGTCATGT GGTGAATATT TGTAAATCAG | 360 |
| TCAGCAGACA GAACGATACT CTTGAAAAAT CTCTTCACAT CATGTCAGCT TCGTCTTTCC | 420 |
| GTATATATGT GACTGACTTC ATCAGTTCTA TCTACAACCT CAAAACAGTG TTTCGAGCTG | 480 |
| ACTTGATCAA TTTTCAAATC TGTACTTTGA GCAAGCTGAG ACTAGCTTCC TATTTGATTT | 540 |
| TCATTGAATA TCAGAAACCC ATTCTCCATC AAATAATTG ACTGCGTCTA ATAATTTTGG | 600 |
| ATCTGGCACG GTGTCTGAAA TAAAGGTGTG GTATTTGGAG AGGGGATTAA TTTTAAAAAA | 660 |
| TCCAGTCTTG TAAAAATTAG AACTATCAAT CAGTAAGATG GTTTCATGGG CTTTGTCAAT | 720 |
| AATATTCCTT TTTGAAATAG CTTGGCTGAG AGAAGCTTCA TAAACATATT GGTCATCAAT | 780 |
| ACCTCTTGCT GAACAAAATG CTAAATCGAT ATTAAATGA TCTAATAAAG AATTTTCCTT | 840 |
| ATCATAGTTG ACCACGGAAC AGGATTGATG TTTGACCTCG CCAGATGTGA TAAAGATTTT | 900 |
| GGAGCTATCT TTAACAGTTT CAGATAGGGT TTGTGCAGTA TGTAACCATT TTGTAAAAAT | 960 |
| AATCAAATTA TCAAGTTCAG AAAGATAGGG ACAGAGTTGC TAGACAGTAG TACTAGAATC | 1020 |
| TAGATAGATA CACATACCAG ACCGAATAAA GTCTTTAGCG AGACTAGCGA TTAGTCTTTT | 1080 |
| TTGCCTAGTA CTTTCTCCTT CACGTATTTG ATGAGAAAGT TCAATTGTGT TCATAGAGGA | 1140 |
| CAGGCTCAGC TATCCGTGCT TTCTTTTGAT AAGACCTTGA TTTTCTAAGA AAATTAAATC | 1200 |
| ACGACGTAAG GTACTTGTGC TGGAGAAAGT GATTTCTGCC AGCTCTTTTA CGGCAATTCT | 1260 |
| TTTTTTCTTT TTGATAATTT CAATCAATTC AAGTACACGT TCATCTTTTA TCATAAGCTC | 1320 |
| CTCCTAATTT ATCATTTCAA CTATATTATA GCACAAATTG GAGGAATTTG AATTATTTTT | 1380 |
| ATGAATATTG GGTTAACATT TGAACATTAT TCAAGTAAGC GTTCACATAT TGAAAAAATA | 1440 |
| AAACGTGGGG ATTATAATAA AGTTAATCmA GGACGAAGAG AGAAGAAAAA TGGAAGCGGT | 1500 |
| TTTAGCAATA GATTTAGGTG CGACTTCTGG AAGAGCAATC GTTGGTTACC TTTCTGAAAA | 1560 |
| TAAACTAGTA ATGGAAGAAA TAAATCGCTT TTCTAATCTA CCTATTAGAG TAAAAGGGCA | 1620 |
| TTTATCTTGG GATATTGACT TTCTACTAGC TAAAATCTT GAAAGTATCC GCTTGGCTAA | 1680 |
| TACTAGTTAC AAGATTTTAT CTATCGGTAT TGACACATGG GGAGTTGATT TTGGACTGAT | 1740 |
| TGATAATGAA GGTAAGCTGT TATTACAACC TGTCATTAT CGTGATGAAA GAACAAAGGG | 1800 |
| AGTGTTAAAG GAAATATCTG AAATGACTGA ATTAGAAAAA CTGTATTCAG AGACAGGAAA | 1860 |
| TCAGATTATG GAGATAAATA CCTTGTTCCTA ACTCTTTAAG GCACGTCAAG AATCTCTGTA | 1920 |
| CTCTTTCTAT AAGACCAATA AGATTCTTTT AATGCCAGAT TTGTTTAATT ATCTCTTGAC | 1980 |

644

| | |
|--|------|
| AGGTAAGTTT GCTACAGAAA AAAGCATTGC TTCAACAACCT CAATTATTTG ATCCTAGGAG | 2040 |
| TCAAAATTGG AATCAGAATA TCTTAAACT ATTTGAATTG GATTCATCTT TACTTCCTGA | 2100 |
| AATTGTTTCA GAGGGAATG TTCTTGAAG GATAAAAGAG GAGTATGGTT TAGGCGATAT | 2160 |
| TCCTGTTGTG AATGTTTGTA GTCATGATAC AGCAAGCGCG ATTGTCTCAG TACCTAAGAC | 2220 |
| AGAAGGTAGT TTATTTATTT CATCAGGTAC TTGGTCTTTG GTTGGAGTGG AACTTACTTC | 2280 |
| ACCGATTCTT ACTACCGAAT CCTTCAGTTA TGGATTTACA AATGAAGTCG GTAAAGATGG | 2340 |
| AGTGATTACA TTTCTGAAGA ATTTGTACAGG GTTGTGGATC ATAGAGGAAC TAAGACGTTT | 2400 |
| ATTTGAACGA AGAGGGAAG CCTATCTCTT TGATGATATT AGGACAATGG TGGAGAAAGA | 2460 |
| AAAAGAAAAT CTTCTCTGA TTGATACTGA ATCAACTGAA TTTGCAACAG AATCTGATAT | 2520 |
| GCACAAGACT TTGACAGAAT ATCTAGCTTA TCATCATGAA ACTAGAGACT GGACAGATGG | 2580 |
| ACAACTATTT AAGATTGTTT ATGAAAGCCT AGCTGAAACG TATAGGAAAG CGATAGAGTT | 2640 |
| ACTAGAAGAA CTAACCTATA AGGTTTATAA GAGGATATAT GTGATTGGAG GAGGTGCTAG | 2700 |
| AGCCAGTTAC TTTAACCAAA TGATTGCTGA TAGAACTGGT AAAGAGGTC TTACAGGTTT | 2760 |
| GACTGAGGCT ACAGCTGTGG GGAATATTGT TGTGCAGCTC ATAGCTATGG GACAATTAAA | 2820 |
| AGGGATGGAA GAGGCTCACC ATGTTATTGA GGAGTTCTA CAATTAGAGA GTTATTACTC | 2880 |
| CCAAAAGAAT TAAAAGATT GAGAGTTTGT AAATTTGCCT CCCTCCCCCT TCTTAGCTTT | 2940 |
| TGTGCAGGAA GGGGGGATAA TTGGTGAATT GAAAAATATT TAGTGTTTTG ATATGAGGAG | 3000 |
| GACAAGGATG TCAGATGTAA AACAAGAATT AATTAAATAT GGTAAAGC TAGTAGAAAC | 3060 |
| AGATTTGACG AAAGGAACAG GTGGGAATCT CAGCGTTTC GATCGTGAAG AACAAATGAT | 3120 |
| GGCAATTACC CCGTCGGGTA TTGATTCTT TGAAATCAAA GAATCCGATA TTGTAGTGAT | 3180 |
| GGATATTAAT GGAAATGTTG TAGAGGGAGA ACGCTTGCCA TCTAGCGAAT GGTATATGCA | 3240 |
| TTTGATTCAA TATCAAACTC GTGATGATAT CGATGCAATT ATCCATGCTC ATACAACTTA | 3300 |
| TGCAACAGTA TTAGCTTGTC TCAGAGAACC ACTTCCAGCG AGTCATTATA TGATTGCAGT | 3360 |
| GGCAGGGAAA GATGTTTCGGG TAGCTGAGTA TGCAACATAT GGCACGAAAG AATTGGCTGT | 3420 |
| GAATGCAGCT AAAGCAATGG AAGGTCGTAG AGCAGTTTGA CTAGCGAATC ATGGAATTTT | 3480 |
| AGCAGGTGCA CAAAATTTAT TGAATGCATT TAATATTGTT GAAGAAGTTG AATATTGTGC | 3540 |
| AAAAATTTAT TGTTTAGCTA AGAATTTTGG AGAGCCAGTA GTTCTTCCTG ATGAGGAGAT | 3600 |
| GGAATTGATG GCAGAAAAAT TTAACACATA CGGTCAGAGA AAATAGGGAG GATATTAATG | 3660 |
| TTAAACATA TACCGAAAAA TATTTCTCCA GATTTATTGA AGACTTTAAT GGAAATGGGA | 3720 |
| CATGGAGATG AAATAGTATT AGCTGACGCG AATTATCCTT CTGCCTCATG TGCAATAAG | 3780 |

645

| | |
|--|------|
| CTAATTCGTT GTGATGGTGT AAATATTCCA GAATTATTAG ATTCCATTCT GTATTTAATG | 3840 |
| CCATTAGATA GTTACGTCGA TAGTTCAATT CAGTTTATGA ACGTTGTTTC GGGTGATGAT | 3900 |
| ATTCCCTAAGA TATGGGGTAC CTATAGACAG ATGATTGAAG GTCATGGTAC AGATCTTAAA | 3960 |
| ACGATTACTT ATCTTAGAAG AGAAGACTTT TATGAACGTA GTAAGAAAGC TTATGCTATT | 4020 |
| GTTGCTACAG GAGAACTTC ACTTTATGCT AATATTATCC TTAAGAAAGG AGTAGTTGTT | 4080 |
| GAAAGAGAAA ATGTTCAATA GAGGAATTTT AGTTGCCAGT CATGGTAATT TTGCTAGCGG | 4140 |
| AGCTCTCATG ACCGCAGAAA TGTGTGTTGG TGAGACAACA AATGATAGAG TTAGGACATT | 4200 |
| AGGTTTGATG CCTGGAGAGA ATATTGTAGA GTTTGAGCAT TATTTTAAAA ATCAAGTGGA | 4260 |
| TGAACTGTTA GACTCAAATC AAGAGGTAT CGTTTGTACT GACTTGATTG GAGGAAGTCC | 4320 |
| TAATAATGTG GCTTTGTCAC GGTTTTAAAA TTTGGATTCA GTTGATATTG TAACAGGGTT | 4380 |
| TAATATCCCT CTCCTAGTGG AATTAATATC AAGTTATGAT TCAAAAATCA ATTTAGAAGA | 4440 |
| AATGTGTCAC AATGCTCAAA ATAGTTTGTT TAATGTTAAA CAACAACCTA ACGTAGAGGA | 4500 |
| GGAGAAGAT TTATGTCTAT AGAGTTTGTT CGTATTGATG ACCGTCTGGT ACATGGTCAA | 4560 |
| GTTGTCACTA CGTGGCTAAA AAAGTATGAT ATTGAGCAAG TTATCATTGT TAATGATCGC | 4620 |
| ATCTCAGAAG ATAAAACACG ACAATCTATT TTAAAGATT CTGCACCGGT AGGTTTAAAA | 4680 |
| ATTGTTTTCT TTAGTGTAAG ACGGTTTGTG GAAGTTTAA ACTCTGTGCC AATAAAAAAG | 4740 |
| AGAACAATGC TGATATATAC AAATCCAAAA GATGTGTATG ATTCTATTGA AGGAAATTTA | 4800 |
| AAATTGGAGT ACCTCAATGT AGGACAGATG AGTAAACGG AGGAAAATGA AAAGGTAACG | 4860 |
| GGAGGTGTAG CTCTAGGTGA AGAAGACAAA TATTATTTTA AGAAAATAGT TGATAAGGGA | 4920 |
| ACGAGAGTTG AAATTCAAAT GGTTCCTAAT GATAAAGTTA CAATGTTGGA AAAATTTTTA | 4980 |
| TAAAAATAAT TTAAGGAGGT ACAGTATATG CTATTCACAC AAGCATTACT GGTGACATTA | 5040 |
| GTTGGGATTA TTGCCACTAT TGAATAAAT GGACCGTTAT TTATGATTCA CCGTCCGTTA | 5100 |
| GTTACAAGTG CAATGGTTGG CTTAGTATTA GGAGATTTC CCAAGGTGT TCTTATTGGT | 5160 |
| TCAGCTCTTG AATTAACTTG GCTCGGTGTA ACAGGTATTG GAGGTATAT TCCACCAGAT | 5220 |
| ACTATTTTCA GTGCGATTAT TGGTACTGCA TTTGGTATTT TATCTGGTCA AGGAGAAACT | 5280 |
| GCTGGTATCG CTATAGCAGT TCCAATTGCA GTTGCTACCC AACAGTTGGA TGTCTTGCA | 5340 |
| AAAACCTTAG ATGTTTATTT TGTGAAAAA GCTGATAATG ATGCTAAAA CGGAGATTAT | 5400 |
| TCAAAGATCG GTTTTATCA TTATTCAGT TTGGTTTAA TCACGTTATT TAAAATTGTA | 5460 |
| CCAATTTTCC TAGCTATTAT GCTGGAGGG GAATATGTGG CAGACTTGTT TGCTAAGGTT | 5520 |

646

| | |
|--|------|
| CCACCAATCG TTATGCAGGG ACTTAACTCT GCAGGTGCTT TACTACCTTC AATTGGTTTT | 5580 |
| GGTATGCTTT TAAATATGAT GCTCAAGAAA AATATGTGGG TATTCTTGTT GATTGGATTTC | 5640 |
| ATTTGTCTCG TGTATGGAGG AATGTCAACC ATTGGGATCT CACTAGTTGG TATTGCGGTA | 5700 |
| GCATACTTCT ACGATATGAT TGAAGCAAA CCACAAGAAA CAACTTCAAG TAGTGATGTT | 5760 |
| GAGGAGGATC TTGATCTATG ATGAATAATA AAGTAACTAA AGTTGAACTT AAAAAAGTTT | 5820 |
| TCAAACGAAG TTTTATGTAT GGTCTTCAT GGAACATGA GAGAATGCAG AACCTAGGTT | 5880 |
| TTCTATATAC AATTCTTCCA GTATTGAAAA AACTATACCC AGACAAAGAT TCAGCTTCTC | 5940 |
| CTGCAATGAA ACGTCACCTT GAGTTTTTCA ATACTCATCA AACAGCGGCA CCATTTATTC | 6000 |
| TTGGAGTTAC TTCCGCTATG GAAGAACAAG AAGGAAATGA AGGTGCAGCT TCAATTACTG | 6060 |
| GTATTAAAGT TGGCTTGATG GGGCCACTGG CTGGTCTAGG AGATAGTTTG TTCTGGCTGA | 6120 |
| CACTAGTTCC TATCTGTTTT AGTATTGGTG CGTCTTATTC TAAAGACGGC GGTGCTTTAG | 6180 |
| GTATCTTTAT CGCCTTAATA TTGTTTAATA TTATTAATAT TCCTGTAAA TATTTGCGTT | 6240 |
| TGAAATATGG GTATACTAAG GGTCTAGTC TTATCCAAGA AAATAATACA AAAGGAACAT | 6300 |
| TGAATCGCGT TACGAGTATG GCGACAGCAT TAGGGCTAGT ACTAGTGGGT GGTTTGATTC | 6360 |
| CATCAATGGT TGGTATTAAT TTTGGATTAG AATTTAAGCA GGGGGAACCT GTTATTTCTG | 6420 |
| TTCAAGAAAT GATTACAAAA TTAATTCCAG GATTTATCCC TATGGCTTTG ACTTTATTAA | 6480 |
| TGTGTAAATT AATTAGAAAA GGAAGAATC CGGTGTACT AATCTTTAGT GTTATGGCTA | 6540 |
| TTGGAGTTAT TCTAGTTGTT TTAGGAATTT TGAAGTAGTA GAAAGTGTGG AGGTGGTATT | 6600 |
| TGGGATATCA CCTCCATTTT GGAAGAGAGG TAAAGAGTGA AATTATGGTA TAAGAAAGCT | 6660 |
| GCCGCAAATT GGAATGAAGC CTTGCCGATT GGAACGGTC ATTTAGGTGG TATGATTTAT | 6720 |
| GGTTCAGCTA CAAAAGAATG TATTCAACTA AACGATGAGA CTATTTGGTA TAGAGGAAAG | 6780 |
| TCAGATAGAA ATAATCCAGA CTCACTATTG CATCTTAAAA AAATTCGGGA ATATCTTTTA | 6840 |
| GATGGAGAAA TTCAGAAAGC CGAAGAATTG ATAAAGTTAA CAGTGTTTGC TACCCCAAGA | 6900 |
| GATCAAAGCC ACTATGAATT ACTTGGGGAA CTTTACATTG AGCATATAGA TATTCAGTCT | 6960 |
| TGTGCTCTTT CATTGTATGA AAGAGAGCTA GATTTAGATA CAGCTATTTT TAATGTTGTG | 7020 |
| TTTGAGCCTA ATAGTTGTAA TTTACAAATA AAAAGAGAAT ATTTTACGAG TTTTAATAAG | 7080 |
| AATATTTTAT GTTGCCGTAT AGTGTCATCA GTTCAAAACA CATTAATTTT AAACATTAAT | 7140 |
| TTGGGTAGAA ATAAACGGTT TAATGACGAA GTATCTAAAC TGGATTCAAG TACAATTTTA | 7200 |
| ATGTCGGCCT CTGCTGGAGG TAGAAAAGGT GTTCAGTTTA AAGTAGTATG TCATCTAAG | 7260 |
| GTTACGGATG GTGAAGTAAG TGTATTGGGA GAGACAATAG TTATTCGGAA TGCTACAGAG | 7320 |

647

| | |
|--|------|
| GTATTTCTTT ATCTCAAATC AATGACGGAT TATTGGGGAA ATATAGATAT TTCTTCTCTT | 7380 |
| CAGGGAGAAT TTAGTAGTAT TGATTACTTT ACAGAAAAAG ATGAACATGT AAAAAAATAT | 7440 |
| CAGGAGCAAT TTAATAGAGT TGATTTTAAA CTAGACTATA GTAAAGGTTG TCTTAGCATT | 7500 |
| CCAACGAATC TACTTCTTGA AAACACTAAA AAGTATAGTA ACTACTTGAC TAACTTGTTA | 7560 |
| TTTCATTATG GAAGATATCT GTTAATATCG TCTAGTCAAC CGAATGGTTT ACCTGCCAAT | 7620 |
| CTTCAAGGAA TATGGTGTGA TGAATTAAAT CCAATTTGGG GTTCTAAATA TACGATTAAT | 7680 |
| ATTAATACTC AAATGAATTA TTGGATGGTA GGTCCATGTG ATTTACCAGA AGTAGAATAT | 7740 |
| CCATTATTG ATATGCTCGA AAGAATGAGA GAACCGGGAA GACTAACCGC TAAGAAAATG | 7800 |
| TATGGAGCTA GAGGTTTAC AGCACATCAT AATACGGATG GTTTTGGCGA TACGGCTCCC | 7860 |
| CAATCTCATG CCATGGGGGC TGCAATTTGG GTATTAAC TAATCCATGGTT ATGTACTCAT | 7920 |
| ATTTGGGAAC ACTATTTATA TTTCCAAGAT GAGCGTATTC TTACGGAACA TTTTGAAATG | 7980 |
| ATAAAGAAG CATTTCTTTT CTTTGAAGAT TATTTATTG AGGTGGATGG CTACTTGATG | 8040 |
| ACAGGTCCAA GTGTCTCACC GGAAAATAAA TATCGCTTAA AAAATGGTAT TGAAGGAAAT | 8100 |
| GCTTGTCTAT CATCTACAAT TGATAATCAA ATTCTAAGAT ATTTTGTGA TTCATGCATT | 8160 |
| GGCATTGCAA AACAATTAGG AGACAATTCG GATTTTATTA GTCGTGTGAA GGAGTTAAAA | 8220 |
| AAGAACTAC CTAAAACAAA AATAGGTAGT AATGGGCAA TCCAAGAATG GTTAGAAGAT | 8280 |
| TATGAAGAAG TAGAGCCTGG GCATAGACAC ATTTACCTC TATTTGGGCT TTATCCTTAT | 8340 |
| AATGAGATTG ATATTCATAA AACTCCGAA TTAGCAGAAG CAGCTAAAAT CACTATCAAT | 8400 |
| AGGAGATTAT CAAACGCTAA TTTTATCT TCACAGGAGA GGGAGCAAGC GATTAATAAT | 8460 |
| TGGTTAGTAA GTGGTTTGCA TGCTAGTACA CAAACAGGTT GGAGTGCTGC ATGGCTGATT | 8520 |
| CATTTTTTTG CGAGACTATA TCAAGGTGAA CCTGCTTATA ACCAGATTAA TGGTTGTTA | 8580 |
| AATAATGCGA CTCTGGCAA TTTATTCTT GACCATCCAC CATTTCAAAT TGATGGTAAT | 8640 |
| TTAGGTTTGG TGAGTGGAAT TTGTGAATTA TTAGTACAGA GCCATCATAA TTGGTTATCA | 8700 |
| CTAATTCCAG CTTTACCTTC TGCTTGGTCA GAAGGAGAAG TGAAAGGTTT CAGAGTAAGA | 8760 |
| GGAGGATATA AGGTATCGTT TGCTTGGAAA AATGGGGATA TAACATTCCT AAAATTGGAA | 8820 |
| GGAGGAAACA AAGATCAAAA AGTAAGAGTA AGAATATATG GCAAAAATAC TGATGTACAA | 8880 |
| AATATTGAAT TGGTATTTAA TTCAGAAAAA ATTATTGAGT TAAATTTTA GGTATAAGTC | 8940 |
| ATGAATAAAG AAAAAATAAA AAGAAAAATTA ATCACAATAT TGTTTGTATG TATTGGGATG | 9000 |
| TTATGTTTTG GATTGTTAGC AGGAGTTAAG GCTGATAATC GTGTTCAAAT GAGAACGACC | 9060 |

648

| | |
|--|-------|
| ATTAATAATG AATCGCCATT GTTGCTTTCT CCGTTGTATG GCAATGATAA TGGTAACGGA | 9120 |
| TTATGGTGGG GGAACACATT GAAGGGAGCA TGGGAAGCTA TTCCTGAAGA TGTAAGCCCA | 9180 |
| TATGCAGCGA TTGAACCTCA TCCTGCAAAA GTCTGTAAAC CAACAAGTTG TATTCCACGA | 9240 |
| GATACGAAAG AATTGAGAGA ATGGTATGTC AAGATGTTGG AGGAAGCTCA AAGTCTAAAC | 9300 |
| ATTCCAGTTT TCTTGTTAT TATGTCGGCT GGAGAGCGTA ATACAGTTCC TCCAGAGTGG | 9360 |
| TTAGATGAAC AATTCCAAAA GTATAGTGTG TTAAAAGGTG TTTTAAATAT TGAGAATTAT | 9420 |
| TGGATTTACA ATAACCAGTT AGCTCCGCAT AGTGCTAAAT ATTTGGAAGT TTGTGCCAAA | 9480 |
| TATGGAGCGC ATTTTATCTG GCATGATCAT GAAAAATGGT TCTGGGAAAC TATTATGAAT | 9540 |
| GATCCGACAT TCTTTGAAGC GAGTCAAAAA TATCATAAAA ATTTGGTGTG GGCAACTAAA | 9600 |
| AATACGCCAA TAAGAGATGA TGCGGGTACA GATTCTATCG TTAGTGGATT TTGGTTGAGT | 9660 |
| GGCTTATGTG ATAAGTGGG CTCATCAACA GATACATGGA AATGGTGGGA AAAACATTAT | 9720 |
| ACAAACACAT TTGAAACTGG AAGAGCTAGG GATATGAGAT CCTATGCATC GGAACCAGAA | 9780 |
| TCAATGATTG CTATGGAAAT GATGAATGTA TATACTGGG GAGGCACAGT TTATAATTTC | 9840 |
| GAATGTGCCG CGTATACATT TATGACAAAT GATGTACCAA CTCCAGCATT TACTAAAGGT | 9900 |
| ATTATTCTTT TCTTTAGACA TGCTATACAA AATCCAGCTC CAAGTAAGGA AGAAGTTGTA | 9960 |
| AATAGAACAA AAGCTGTATT TTGGAATGGA GAAGTAGGA TTAGTTCATT AAACGGATT | 10020 |
| TATCAAGGAC TTTATTCGAA TGATGAAACA ATGCCCTTAT ATAATAATGG GAGATATCAT | 10080 |
| ATTCTTCCTG TAATACATGA GAAAATTGAT AAGGAAAAGA TTTCATCTAT ATTCCCTAAT | 10140 |
| GCAAAAATTT TGACTAAAA TAGTGAGGAA TTGTCTAGTA AAGTCAACTA TTTAAACTCG | 10200 |
| CTTTATCCAA AACTTTATGA AGGAGATGGG TATGCTCAGC GTGTAGGTAA TTCTGGTAT | 10260 |
| ATTTATAATA GTAATGCTAA TATCAATAAA AATCAGCAAG TAATGTTGCC TATGTATACT | 10320 |
| AATAATACAA AGTCGTTATC GTTAGATTTG ACGCCACATA CTTACGCTGT TGTTAAAGAA | 10380 |
| AATCCAAATA ATTTACATAT TTTATTGAAT AATTACAGGA CAGATAAGAC AGCTATGTGG | 10440 |
| GCATTATCAG GAAATTTTGA TGCATCAAAA AGTTGGAAGA AAGAAGAATT AGAGTTAGCG | 10500 |
| AACTGGATAA GCAAAAATTA TTCCATCAAT CCTGTAGATA ATGACTTTAG GACAACAACA | 10560 |
| CTTACATTAA AAGGGCATAC TGGTCATAAA CCTCAGATAA ATATAAGTGG CGATAAAAAAT | 10620 |
| CATTATACTT ATACAGAAAA TTGGGATGAG AATACCCATG TTTATACCAT TACGGTTAAT | 10680 |
| CATAATGGAA TGGTAGAGAT GTCTATAAAT ACTGAGGGGA CAGGTCCAGT CTCTTTCCCA | 10740 |
| ACACCAGATA AATTTAATGA TGTAATTTG AATATAGCAT ATGCAAAACC AACAACACAA | 10800 |
| AGTTCTGTAG ATTACAATGG AGACCCTAAT AGAGCTGTGG ATGGTAACAG AAATGGTAAT | 10860 |

649

| | |
|--|-------|
| TTTAACTCTG GTTCGGTAAC ACACACTAGG GCAGATAATC CCTCTTGGTG GGAAGTCGAT | 10920 |
| TTGAAAAAAA TGGATAAAGT TGGGCTTGTT AAAATTTATA ATCGCACAGA TGCTGAGACT | 10980 |
| CAACGTCTAT CTAATTTTGA TGTGATTCTA TATGACAATA ATAGAAACGA AGTTGCTAAG | 11040 |
| AAACATGTTA ATAATTTGTC GGGTGAATCT GTTAGTCTAG ATTTCAAAGA AAAAGGAGCA | 11100 |
| AGGTATATTA AAGTTAAATT ACTAACGAGT GGAGTGCCTT TGAGTTTAGC AGAAGTAGAG | 11160 |
| GTTTTTAGAG AATCAGATGG TAAGCAATCT GAAGAGGATA TAGATAAAAT AACAGAAGAT | 11220 |
| AAAGTAGTCT CTACAAATAA GGTAGCTACT CAAAGTTCAA CCAATTATGA GGGTGTAGCT | 11280 |
| GCTTTAGCAG TTGATGGTAA TAAAGATGGA GATTACGGAC ATCATTCCGGT GACTCATACT | 11340 |
| AAGGCAGATT CTAACGCTTG GTGGCAGGTC GATCTGGGAG AAGAGTTTAC GGTTCCTAAA | 11400 |
| GTGATATTT ATAATAGAAC AGATGCCGAA CCTCAGCGTT TATCTAATTT TGATGTTATT | 11460 |
| TTTCTATCTT CATCAGGAGA AGAAGTTTTT AGAAGACATT TTGATAAAGT AGTTGATGGT | 11520 |
| TTGTTATCTT TAAAAGTACC TTCTGTAGGG GCTAAGCTAG TCAAAATAGA ATTAAAATCA | 11580 |
| GCAGCTATTC CGTTAAGTTT AGCGGAAGTT GAAGTCTATG GTTCAAAGAG AACTCCGAAG | 11640 |
| AACTTTCTA ATATTGCATT AACAAAAGAA ACTCGACAGA GTTCAACGGA TTACAATGGT | 11700 |
| TTTTCTCGTC TAGCAGTTGA TGGAAATAAA AACGGAGATT ATGGTCATCA TTCAGTGACT | 11760 |
| CATACCAAAG AAGATTCTCC TTCATGGTGG GAGATAGATT TAGCACAAAC CGAAGAATTA | 11820 |
| GAAAAGTTAA TTATTTATAA TAGAACAGAT GCTGAAATTC AGAGATTATC AAATTTTGAT | 11880 |
| ATTATTATAT ATGATTCAAA TGATTATGAA GTTTTACAC AACATATTGA CAGTTTAGAA | 11940 |
| AGCAATAATC TATCCATAGA CTTAAAAGGA CTGAAGGGAA AAAAGGTTAG AATTTCTTTG | 12000 |
| AGAAGCGCAG GAATTCCTTT AAGTTTAGCA GAGGTAGAGG TTTATACTTA TAAGTAATTT | 12060 |
| TAAAAATTAT CACCCAGGCT ACCGTAAATA TAATGGAGAT GGTAGTATGA AAGAAACAGA | 12120 |
| AAAATAAGAG GAAAATAGTA TGATTCAACA TCCACGTATT GGGATTCTGC CGACTATTGA | 12180 |
| TGGTCGTCGT CAAGGTGTAC GCGAATCACT TGAAGTGCAA ACAATGAACA TGGCTAAAAG | 12240 |
| TGTGGCAGAT TTGATTTCAA GCACATTGAA ATATCCAGAT GGGGAACCTG TGGAATGCGT | 12300 |
| GATTTCTCCA TCTACTATTG GCCGTGTACC AGAGGCTGCA GCTTCCCATG AGTTGTTTAA | 12360 |
| AAAATCAAAT GTTTCGCGAA CAATTACAGT TACACCATGC TGGTGTTATG GTAGTGAAAC | 12420 |
| TATGGATATG TCTCCAGATA TTCCTCATGC TATTTGGGGA TTTAATGGGA CAGAACGCCC | 12480 |
| AGGAGCTGTC TATCTTGCAG CTGTACTAGC TTCACATGCT CAAAAAGGGA TTCCAGCCTT | 12540 |
| TGGGATTTAT GGAAGAGATG TTCAGGAAGC TAGTGACACA GATATTCCAG AAGATGTCAA | 12600 |

650

| | |
|--|-------|
| AGAAAACTT TTACGCTATG CGCGTGCAGC TCTTGCAACT GGCTTGATGA GAGACACTGC | 12660 |
| TTACCTATCA ATGGGTAGTG TTTCGATGGG GATTGGTGGT TCTATTGTAA ATCCGGATTT | 12720 |
| CTTCCAAGAA TACTTAGGAA TGCGAAATGA ATCGGTAGAT ATGACGGAGT TCACGCGCCG | 12780 |
| TATGGACCGT GGTATTACG ACCCTGAAGA GTTCGAACGT GCGCTCAAAT GGGTGAAAGA | 12840 |
| AAACGTAAAA GAAGGATTCG ACCATAACCG TGAAGACCTT GTTTTAAGCC GTGAAGAAAA | 12900 |
| AGATAGACAA TGGGAATTTG TTATTAAGAT GTTCATGATT GGACGTGACT TAATGGTTGG | 12960 |
| TAACCCAAGA CTGCTGAAC TTGGTTTGA GGAAGAAGCG GTTGGTCACC ATGCTTTAGT | 13020 |
| AGCTGGTTTC CAAGGTCAAC GTCAGTGGAC AGACCATTTT CCAAATGGGG ACTTTATGGA | 13080 |
| AACTTTCCTC AATACTCAGT TTGACTGGAA TGGTATTCGA AAACCATTTG TATTTGCGAC | 13140 |
| AGAGAATGAT TCACTAAATG GTGTGTCTAT GCTCTTTAAT TATCTATTAA CAAATACTCC | 13200 |
| ACAAATCTTT GCTGATGTGC GTACTTATTG GAGCCCAGAG GCTGTTAAAC GTGTAACGGG | 13260 |
| ACATACTTTA GAGGGTCGTG CTGCAGCTGG CTTCTTACAT CTAATCAACT CTGGTTCTTG | 13320 |
| TACATTTGGT GGTACAGGTC AAGCTACTCG AGATGGCAAA CCTATTATGA AACCATTTCTG | 13380 |
| GGAGTTGGAA GAAAGTGAAG TGCAGGCTAT GCTTGAAAT ACAGACTTCC CACCAGCAAA | 13440 |
| CCGCGAATAC TTCCGTGGAG GAGGATTCTC AACTCGTTTC TTGACGAAGG GGGATATGCC | 13500 |
| AGTAACAATG GTACGTCTCA ATCTTCTAAA AGGGGTGGT CCAGTGCTAC AAATTGCAGA | 13560 |
| AGGTACACA CTTGAACCTC CTGAAGATGT TCACCATACT TTAGATAATC GTACAGATCC | 13620 |
| AGGATGGCCA ACTACTTGGT TTGCTCCACG TTGACAGGA AAAGGTGCTT TCAAGTCTGT | 13680 |
| CTATGACGTC ATGAATAATT GGGGAGCTAA TCACGGAGCC ATAACATATG GACACATTGG | 13740 |
| AGCAGACTTG ATTACCTTGG CTTCTATGTT GAGAATTCCT GTCAATATGC ATAATGTACC | 13800 |
| TGAGGAAGAT ATCTTTAGAC CTAAAAATG GTCCTTATTT GGAACAGAAG ATCTAGAATC | 13860 |
| AGCAGACTAT CGTGCATGTC AGTTGTTGGG GCCACTACAT AAATAAACT TGTTTATATA | 13920 |
| GGAGGTGAAC TTACGTCCCT CCTATCCTTT TAAAAAGATT TGTAAACAA TTCACAAATA | 13980 |
| ATTGAAAACG AATACAAAAA GTAATATAAT GATGTTAAAT AGATAGCGCG GAGGCGCAGG | 14040 |
| AGGAAAATTA TATGGCTATA TTTTATGTTT CGGCAGTCAA CCTTATTGGA AAAGGTGTTG | 14100 |
| TAAATGAAGT GGGTCCTTAT ATCAAGGAAC TTGGCTATAA AAAGGCACTT TTGGTGACAG | 14160 |
| ATAAGTACAT CGAAGGCAGT GATATTTTAC CTAAGACTTT AAAACCACTG GATACAGAAG | 14220 |
| GAATCGAATA T | 14231 |

(2) INFORMATION FOR SEQ ID NO: 82:

(i) SEQUENCE CHARACTERISTICS:

651

(A) LENGTH: 16995 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 82:

| | |
|--|------|
| AGTTCTCTTA ACTTTTCTAG GATGGCATTC TCCGCTCTCA GGTACTCATT TTCTGCTgAA | 60 |
| GACGTTCTAA TTCTGTCTC TCTTCAGGTC TCGTTTTTGG CTTACGTCCC ATTTTAGGTA | 120 |
| CTCTCCCTCT TGTTTTCTCA ACAATAGTAT ACCCGTTTTT CCGTATTGT GCTAGCCAGT | 180 |
| TAAGAAGTAT CGTACGACTT GGGAGACCGT ATTCAAGAGA AACTCTATCT TTAGTCCAGC | 240 |
| CTTCATGTCA GACTTTATTA CTCATTTCTT GTTTTAAATC AGGAGAATAG TAACGATTTT | 300 |
| TTCCTTTTTT GACGAACTCT ATTCCGTAAC GATCAATCAA TTAAATCATG TACCTAATAT | 360 |
| TAGAATGCT TATCCCAAAT TTATTTGAAA GCTTCTCTAA GCTATATCCT TGTTTTCTAA | 420 |
| GTTCATAGAT CTGAACTTTA TCATCATAAG TTAGTTTCAT AATAAAAACA CCCCAAAGT | 480 |
| TAGATTTTTT CTGTCTAACT TTTGGGGTGT AGTTCATGTA CACCTGATAT GATGCGTTTT | 540 |
| ATAATTTTTA AGCCTTTTTG CCCAGCCTCG TCAAAAGTAA TGTTTTGACA CAAAATCTGT | 600 |
| GACAAAACCT TAGTTTAAAG GGTTTTTAAC TTTGTATATA CTAGTTTAA GAAAAGGAGG | 660 |
| ATGATCTAAT GGAAGAAAA GTATCATTGA AAGTCAGGGT TCAAAAAC TA GGGACATCGC | 720 |
| TTTCAAATAT GGTATGCCC AATATTGGAG CATTTATTGC TTGGGGAGTA TTGACTGCCC | 780 |
| TCTTTATCGC TGATGGCTAT CTGCCAAATG AACAGTTAGC TACTGTGTGT GGTCTATGT | 840 |
| TAACGTATTT ATTGCCAATC CTGATTGGTT ACACAGGTGG ATATATGATC CATGGCCAAC | 900 |
| GTGGTGCCGT TGTAGGAGCT ATTGCTACTG TTGGTGCAAT CACAGGTTCT AGTGTTCCTA | 960 |
| TGTTTATCGG AGCTATGGTA ATGGGCCCAC TGGGAGGATG GACTATCAAG AAATTTGATG | 1020 |
| AGAAGTTCCA GGAATAAAT CGTCCCGGAT TTGAAATGTT AGTTAATAAC TTCTCAGCTG | 1080 |
| GTCTCGTTGG TTTTGCAATTA TTGCTTTTGG CTTTCTACGC AATCGGTCCA GTCGTATCGA | 1140 |
| CTCTTACTGG AGCTGTGGG AATGGTGTG AGGCTATTGT CAATGCTCGC CTCCTTCCTA | 1200 |
| TGGCTAATAT TATCATCGAA CCGGCTAAAG TCCTTTTCCT CAATAATGCC CTCAATCATG | 1260 |
| GCATTTTAC TCCTCTGGGA GTAGAACAGG TAGCTCAAGC TGGTAAGTCA ATTCTCTTCC | 1320 |
| TATTGGAAGC TAATCCTGGA CCAGGTCTGG GAATCTATT AGCTTATGCT GTATTGCGTA | 1380 |
| AAGGTCTGTC TAAATCTTCT TCTTGGGGG CAATGGTTAT TCATTTCTTC GGAGGGATTC | 1440 |
| ATGAAATTTA CTTTCCTTAT GTTATGATGA AGCCTACTCT ATTTTCTAGCT GCTATGGCAG | 1500 |

652

| | |
|--|------|
| GAGGTATCTC TGGAACCTTT ACTTTTCAAC TCTTAGACGC TGGTCTTAAA TCTCCAGCTT | 1560 |
| CACCAGGTTC TATTATTGCG ATTATAGCTA CGGCGCCAAA AGGTGTTTGG CCCCATCTAA | 1620 |
| ATGTTCTTTT AGGTGTTTFA GTGGCAGCAG TTGTTTCTTT CCTTGTAGCA GCCCTTATTC | 1680 |
| TTTCATGCAGA CAAGTCAACT GAGGATTCGC TCGAAGCTGC TCAGGCGGCT ACCCAAGCAG | 1740 |
| CTAAGGCTCA GTCTAAAGGT CAGTTAGTAT CAACTTCTGT TGATGCAGTT GTTTCGACAG | 1800 |
| ACTCAGTGGA AAAAATCATT TTCGCCTGCG ATGCTGGTAT GGAAGCTCT GCTATGGGAG | 1860 |
| CTAGTATTCT TCGAGATAAG GTTAAAAAAG CAGGTCTAGA GATTCCAGTA TCTAATCAGG | 1920 |
| CAATCTCAAA TTTGCTTGAT ACACCAAAAA CATTAATTGT TACTCAGGAA GAACTGACAC | 1980 |
| CAAGAGCTAA AGACAAGAGT CCAAGTGCTA TTCATGTTTC TGTGATAAT TTCTTAGCGT | 2040 |
| CCTCTCGTTA TGATGAAATT GTAGCTTCAT TAACAGGAGC TTCTCCAATA GCAGAAATTG | 2100 |
| AAGGAGATAT ACCAACTTCA GCACCAGTAG ATAGTCAGGA AAGTGACCTT AACCATATTG | 2160 |
| ATGCTGTAGT AGTTGCTTAT GGTAAAGCAC AGGGAAGTGC AACTATGGGC TGTGAAACGA | 2220 |
| TTGCGGCTAT TTTTAGAAAC AAGAATATTC GTATTCCAGT TTCTACTGCC AAAATTTTCAG | 2280 |
| AATTAGGTGA ATTTAATTCT AAAACATAA TGATTGTAAC AACTATTTCT TTACAGGCAG | 2340 |
| AAGTGCAGCA AGCAGCACCG AATTCTCAAT TTCTTATTGT GGATAGTTTA GTAACAACAC | 2400 |
| CAGAAATGA CAAAATGGCT GCTAGAATGT ACAAATAGAA CTAGAGGTTT CTAAATTACG | 2460 |
| AATGCTATTA ACCAAACGAG AAGAACAATT ATTGAAGGCT TTCCTACATG TAGGGAAGCT | 2520 |
| TTCAATGCAA GATATGACTG AAATCTTACA GGTTCATCT AGAACAATTT ATCGAACTTT | 2580 |
| ATCAGATTTG ACAGATAGCA TGGAGCAATA TGAATCGAA ATAACGAAGC ATGGGAAATA | 2640 |
| CTATATTTTG ACTGGAGAGT TGGATGATT GCCGACAGAA CTGGAAGTGT TAGTTGAGTA | 2700 |
| TAGTCCCAA GAAAGACAAG AGTTGATTAC CTATCGCCTT CTGACTGAGA GTGGTTTTGT | 2760 |
| CACCAATGAA GCATTGCAAG AGTGCACGAA AGTCAGTAAT GTAACATTA TTCAGGATAT | 2820 |
| TTTCAGATATT GATAAGCGTC TTTTAGACTT TGATCTGAAA ATTGAACGAC AAAAAGGTTA | 2880 |
| TCGGATTCTT GGTGATTCAG TTGGTAAGAG AAGATTTTGT GCTATTTTAC TGACAAACTG | 2940 |
| TATCTCAGTA GCAGATTTTT CAACCGGTAA TTTTGGGAGC TTTGATATTT TAGAAGCAGA | 3000 |
| TAGAACTGGG CTGGCCAGTC AGATTGTTAA TAAGCAACTG TCAGGTTTTT CAGATATGGA | 3060 |
| TGCTAGGATG AAGATGTTTT TTGCGATCTT GTTATCTCT ATAGGTCAGG AGCAAAACAT | 3120 |
| TGAAAATTCA CCTAATACTA GTAAGCAGGC TTTGGAAATT TCTCAAAAAA TTTTCAAGC | 3180 |
| TTACTCTAAG CAGACTGCAC AATTTTATAG TATTCAGGAA ATTATCTATT TTGCGAGCAT | 3240 |
| CTTGGATGAA TTAATCATTA AACGTCAGGA CAATCCGCTC TTTACGGAGA AATTTGATGG | 3300 |

653

| | |
|--|------|
| TGAATTTTTC TACAATATTT CAAATCTGAT TGATACGGTT TCCATGTATA CCAAGATTGA | 3360 |
| CTTTTTTAAG GACAAGGTTT TATTCATTTT TCTTTTCCAT CATATTCGGC TCAGTTTAGG | 3420 |
| CGTCCCTATC CTTTTCAGG GTGAAAATTT GCCAGAATCT ATCCAGATTT TAGTTGAAAG | 3480 |
| GAATAAATTT CTTTATACAG TCATCAGTCT TTTAGTGAAT GATATTTTTC CGAAATATCT | 3540 |
| TCATACAGAG TATGAGTATG GCATGATTGC CCTACATTTT ATCTCTAGCT TAGGCCGTAG | 3600 |
| .TCCAGAGATT TATCCAGTCC GTGTTTGCT TTAAACGGAT GAACGTCGGG TCACTAGAGA | 3660 |
| TTTATTAGTC AGTAAATTA AGAGTGTTC TCCTTTTGTA GAGTTGATAG ATATTCAATC | 3720 |
| TCTAGTAGAT TACCACAGTA TTGATCTCAG TCAGTATGAT TATATTTTAT CTACCAAGCC | 3780 |
| GCTGACTAAT CAGGAAATCG ATGTAATTTT TAGTTTCCA ACCGTCAAAG AATTGCTTGA | 3840 |
| ATTACAGGAA CGACTTCAGT ATGTACAGGC ACATCGTACA ATTGTCGCGC GTGATGCTAT | 3900 |
| CGCTCCAGAG AAAAGTTATG ACTTGCAAGA TTATTTAATA TCTAGTAGTC AGCTTTTGAG | 3960 |
| TCAATTCGAG TTGGTTCAAT TGGAGAATAA TCAATCATTT GAGCACACGG TAGAACAAAT | 4020 |
| CATCCAATAT CAGAAGAATG TGAGTGACAG AGCTTACCTA ACAAGAAAAT TGTATCTCA | 4080 |
| CTTCCAGAAT AGTCCTATGG CTATTCCTAA TACTGGTCTG GTGCTTTTAC ATAGTCAGTC | 4140 |
| TAGCAAAGTA ACAACAAATA GTTTTACTAT GTTTGAACTC AAACCTACCTA TCTCCGCATT | 4200 |
| GTCAATGAAA CGAGAGGAAG AAGAGGTCAA AAGGTGTCTG CTAATGCTAA TGTCTAAAGA | 4260 |
| AGCTAGCGAG GAAGCTAGAG ATTTAATGAC AGCTATTAGT CAGTCGATTA TTGAAAATCA | 4320 |
| TCTTTATACA GAGATTTACA AGACGGGAAA TCAATCCATT ATTTATCAGA TGCTAAATAC | 4380 |
| TATTTTTAAC GAAAAATTA AGAAATTGGA GAACTAATAT GAACTTGAA AAACATTTGA | 4440 |
| TTAAGCTTAA TAAACAATTT TCTAACAAGG AGGAAGCTAT TTGTTATTGT GGGCAAGTTC | 4500 |
| TTTATGAGGG TGGATATGTT AATGAAGACT ATATTGAAGC CATGATTGAG CGAGATAAAG | 4560 |
| AGCTATCTGT TTACATGGGT AACTTTATCG CCATACCGCA TGGAACAGAT GCAGCAAAAA | 4620 |
| ATGATGTCCT CAAGTCTGGT ATTACAGTCG TTCAAGTCCC TAGAGGGGTT GATTTTGGGA | 4680 |
| ATGTATCTAA CCTCAAGTG GCAACGGTTC TTTTGGTAT TGCTGGTATT GGTAATGAAC | 4740 |
| ACTTAGAAAT TATTCAGAAA ATTTCTATCT TCTGTGCAGA TGTAGATAAT GTTCTTAAAC | 4800 |
| TAGCAGATGC TCAGTCAAAA GAGGAAGTAT TGCCTTATT TGATGCTGTT GAATAATTGA | 4860 |
| ATTTAGTCAT TTGTCATCTA GTATATATGT CCTCAAATA GGAAAAGGAG AAATTGAATG | 4920 |
| AAACATTTCTG TTCATTTTGG TGCCGGTAAT ATCGGTCGTG GTTTTATAGG TGAAATTCTA | 4980 |
| TTTAAAAATG GTTCCATAT TGATTTTGTG GATGTCAATA ATCAGATAAT TCATGCTCTG | 5040 |

654

| | |
|--|------|
| AATGAAAAGG GCAAGTATGA AATTGAAATT GCACAGAAAG GACAGTCTCG TATAGAAGTA | 5100 |
| ACTAATGTGG CTGGCATTAA TAGCAAAGAA CATCCTGAGC AAGTCATTGA AGCGATTCAA | 5160 |
| AAGACGGATA TTATTACTAC TGCAATCGGA CCTAATATAC TCCCTTTTAT CGCCGAACTT | 5220 |
| CTAGCCAAAG GAATCGAAGC TCGCCGAGTT GCAGGAAATA CACAGGCATT GGATGTTATG | 5280 |
| GCCTGTGAAA ATATGATTGG CGGGTCTCAA TTTCTTTATC AAGAAGTCAA GAAATATTTA | 5340 |
| AGTCCGGAAG GTTTGACATT TGCTGATAAC TACATAGGTT TTCCAAATGC TGCAGTAGAC | 5400 |
| AGGATTGTTC CAGCACAAAG TCACGAAGAT TCCCTTTTGT TTGTGGTCGA GCCCTTTAAT | 5460 |
| GAATGGGTCG TGGAAACCAA GCGTCTTAAA AATCCAGATT TACGTCTAAA AGATGTGCAT | 5520 |
| TATGAAGAAG ATTTAGAACC CTTTATTGAG CGAAAACTTT TTTCAGTCAA TTCTGGACAT | 5580 |
| GCAACTTCAG CTTACATTGG TGCGCATTAT GGTGCCAAGA CAATTTTGGA AGCTCTTCAA | 5640 |
| AATCCTAATA TTAAATCTCG GATTGAATCT GTATTAGCTG AAATTCGGAG TCTCTTGATT | 5700 |
| GCCAAATGGA ACTTTGATAA AAAAGAATTG GAGAATTATC ACAAAGTCAT TATAGAACGA | 5760 |
| CTTGAAAACC CTTTCATAGT GGACGAGGTT AGTCGCGTAG CTCGTACTCC AATCCGAAAA | 5820 |
| TTAGGCTATA ATGAACGATT CATCCGGCCG ATACGTGAAT TGAAAGAACT CAGTTTGTC | 5880 |
| TATAAAAACC TACTTAAAC AGTTGGCTAT GTCTTTGACT ATCGCGATGT AAATGATGAA | 5940 |
| GAAAGTATTC GATTAGGTGA ATTGTTGGCT AAACAATCAG TCAAAGATGT TGTATACAA | 6000 |
| GTTACAGGTT TAGACGACCA AGAATTGATT GAGCAAATTG TAGAGTATAT TTAATCTTTT | 6060 |
| TCGAAAATCT CTTCAAATCA GGTTAGCATC GCTTTGTCTT AGGCATATGT TGTCTATCT | 6120 |
| ACAACCTCAA AGCAGTGCTT TGAGCTGACT CCGTCAGTCT TATCTGCAAT CTCAAAACAC | 6180 |
| TGTTTGAGTT ATCTGCGGTA ATCTTTCTAG CTTGTCTTTG ATTTTGTGTTG TTATTTATAA | 6240 |
| GGTAAAAGAA GCTGGACAAA AAGTCTTCAA AATCGGGAAA AGGCAGCCTA TCGGGTGTTT | 6300 |
| AAAAATCTTG ATAGGATGTC CTTTATTATG GAAAGCCTTA TTGGATTTTC TCCTCAGATT | 6360 |
| GAGTTTTTGA TCAGCTTTAT GAGATAGGTC TTGCTAGAGA TGAGCCCAT CATGTTATTT | 6420 |
| TTATGGACAG TGGGAAAATT GTTGAAAAA ATAATGCCCA TCAATTCTTT AGTCGTCCAA | 6480 |
| GAGAAGAACG AACCAAGCAA TTTTGAACG AATTCTTTTCG AATGCGATCT ATATAGTAAA | 6540 |
| ATGAAACAAG AACAGGACAA ATCGATCAGG ACAGTCAAAT CGATTTCTAA AAATGTTTTA | 6600 |
| GAAGTAGAGG TGTACTATTC TAGTTTCAAT CTACTATATA ACTGAAAAAT TAGATAAATT | 6660 |
| AGTTTTGGAA AATGACTAAC CAAAAGATAT CCAAAGTAGT CTAAAATTGT CTATACTTTA | 6720 |
| TGAGTGTTTT AGTTAGGAAA AAGGCTTGTT GTCTATAATT GTCTGCATTA GTCTAGATTT | 6780 |
| TATTTATAGA AAATGTTATA ATAGACTGTA TTTAAAAAT TTTAAGGAGA AATGACAGAA | 6840 |

655

| | |
|---|------|
| TGTCTGTATC ATTTGAAAAC AAAGAAACAA ACCGTGGTGT CTTGACTTTC ACTATCTCTC | 6900 |
| AAGACCAAAT CAAACCAGAA TTGGACCGTG TCTTCAAGTC AGTGAAGAAA TCTCTTAATG | 6960 |
| TTCCAGGTTT CCGTAAAGGT CACCTTCCAC GCCCTATCTT CGACCAAAAA TTGGGTGAAG | 7020 |
| AAGCTCTTTA TCAAGATGCA ATGAACGCAC TTTTGCCAAA CGCTTATGAA GCAGCTGTAA | 7080 |
| AAGAAGCTGG TCTTGAAGTG GTTGCCCAAC CAAAAATTGA CGTAACTTCA ATGGAAAAAG | 7140 |
| GTCAAGACTG GGTATCACT GCTGAAGTCG TTACAAAACC TGAAGTAAAA TTGGGTGACT | 7200 |
| ACAAAAACCT TGAAGTATCA GTTGATGTAG AAAAGAAGT AACTGACGCT GATGTCGAAG | 7260 |
| AGCGTATCGA ACGCGAACGC AACAACCTGG CTGAATTGGT TATCAAGGAA GCTGCTGCTG | 7320 |
| AAAACGGCGA CACTGTTGTG ATCGACTTCG TTGGTTCTAT CGACGGTGTT GAATTTGACG | 7380 |
| GTGGAAAAGG TGAAAACTTC TCACTTGGAC TTGGTTCAGG TCAATTCATC CCTGGTTTCG | 7440 |
| AAGACCAATT GGTAGGTCAC TCAGCTGGCG AAACCGTTGA TGTATTCGTA ACATTCCCAG | 7500 |
| AAGACTACCA AGCAGAAGAC CTTGCAGGTA AAGAAGCTAA ATTCTGACA ACTATCCACG | 7560 |
| AAGTAAAAGC TAAAGAAGTT CCGGCTCTTG ACGATGAACT TGCAAAAGAC ATTGATGAAG | 7620 |
| AAGTTGAAC ACTTGCTGAC TTGAAAGAAA AATACAGCA AGAATTGGCT GCTGCTAAAG | 7680 |
| AAGAAGCTTA CAAAGATGCA GTTGAAGTG CAGCAATTGA TACAGCTGTA GAAAATGCTG | 7740 |
| AAATCGTAGA ACTTCCAGAA GAAATGATCC ATGAAGAAGT TCACCGTTCA GTAAATGAAT | 7800 |
| TCCTTGGA GAA TTTGCAACGT CAAGGGATCA ACCCTGACAT GTACTTCCAA ATCACTGGAA | 7860 |
| CTACTCAAGA AGACCTTCAC AACCAATACC AAGCAGAAGC TGAGTCACGT ACTAAGACTA | 7920 |
| ACCTTGTTAT CGAAGCAGTT GCCAAAGCTG AAGGATTGTA TGCTTCAGAA GAAGAAATCC | 7980 |
| AAAAAGAAGT TGAGCAATTG GCAGCAGACT ACAACATGGA AGTTGCACAA GTTCAAAACT | 8040 |
| TGCTTTCAGC TGACATGTTG AAACATGATA TCACTATCAA AAAAGCTGTT GAATTGATCA | 8100 |
| CAAGCACAGC AACAGTAAAA TAATCTTAAT AAACAGAAAA CCCACCTGAA TTGGTGGGTT | 8160 |
| TTCTGATGCA CTATTTTCCA AAAATCTCTT TGAGGTCTGT GTCTGTAATC CCAATCATGG | 8220 |
| CTGGGATGCG GTCCCAAGTT TCTTCGGTTA GGATGTAGGA TTGTTTCAGAG GCACTTGATG | 8280 |
| TGACTGTTTC AGAGACAGCT TGTGCTTTT CTTCAACATT CTCCAGTAGA TCACTGAAGC | 8340 |
| GTCAATCAG ATAGGTTTTT CGGGCAGTTC CGATGTGTTG GGTAGCATAG TCGAAGGCTT | 8400 |
| GTAATTCGCC TAGTAAGATG AGTTTGCTTT TGGCACGTGT AATGGCTGTG TAGATGAGAT | 8460 |
| TTGCTCCAG CATACGTCGG CTAGCACTAG TAATCGGTAG GATGACAACT GGGAACTCAC | 8520 |
| TTCCCTGAGA CTTATGAATA CTCATGGCAT AGGCCAAGCG AATCTGTAC CATTCGTTAC | 8580 |

656

| | |
|--|-------|
| GGGGGTAAGA GACTTCATTA CCATCAAAAT CAATGACAAT CTCGTCTTGT TTCGATTTCGG | 8640 |
| TGTATTTTACC AGGAATCAGG TCTGTGATAG CTCCTAAATC CCCATTAAAG ACATTGATTT | 8700 |
| CAGCATCGTT AACCAAATGA ATGACCCTGT CTCTCTTACG ATAGTGACAC TGAGGAGCTT | 8760 |
| CAAACTGAG TTGATCTTTT TGTGGGGGAT TGAGCAGGTC TTGCATGAGC TGATTGATAG | 8820 |
| CATCAATCCC TGCCGTCCCT CGGTACATAG GAGCCAGAAC TTGGATATCA CGGGCGGGAA | 8880 |
| TACCATTTCT GAGGGCGGCA CCTAAGATTT TTTCAATGGT GGCAGGAATA TGGCCACTAG | 8940 |
| CAATTTCAAA GTAGGAACGG TCAGCTTTTT TTTGGGTGAA ATCAGCTGGC AAGATGCCCT | 9000 |
| GTCGAATCTG ACTAGCTAGG GTGACGATGG TTGATTCCTT GCTTTGTCTG TAAATTTTTT | 9060 |
| CCAAGCGAGT CTGAGGAATC AAAGGAATAT GAAGTAGATC CGCTAGAACC TGTCACGGAC | 9120 |
| TGACAGAAGG TAGCTGATCA CTGTCACCTA CGATGAGGAT CTTACTGTTA GAAGAGATAT | 9180 |
| TGGAGAAGAG TTGATTGGCC AGCCAAGTAT CTACCATAGA GAATTCATCC ACGATGATAA | 9240 |
| AGTCAGCATC TAGGTAATCT TCCAGATGAC TGGTATCATC GTCACCTGTC ATTCCCAAGT | 9300 |
| GGCGATGTAT GGTGCGGCTA GGCAACCTG TCAATTCATT CATGCGACGA GCAGCTCGAC | 9360 |
| CAGTTGGAGC AGCAAGAAGA ATGGGCAGAT TGCTTTTCTT CCTGAAGTCA AGTCCTTCTA | 9420 |
| AAAGGGCATA AACAGCAATG ATTCCATTGA TAACAGTTGT CTTACCAGTA CCAGGCCAC | 9480 |
| CTGTCAGGAT AAAGACCTTA TTCTGGATAG CATCACAGAT AGCCTGTMTT TGAATGTTAT | 9540 |
| CATACTCAAT TCCAGTTCT TGCTCGACAG TAGTGATATG TTTTGAATG GTTTCTAAAT | 9600 |
| CATGACTCTT CTGTTTTCCT TTTTCAAGGA TACGAACCA GTGACTGCGG ATGCCTTCCT | 9660 |
| CAGCGAAAAA GAGGCTGTTG TCAAAGATCT TGGTATCAAT CTGCTGAACC TTGTCTTCTT | 9720 |
| CGATCAGGTA GGAGAGCTCT TGGGCAACTT GGCTGGGGTC TAGTTCCACG GGACGGGAAG | 9780 |
| ACTCAAGGAG AGTAAGGGTT TGTTCAGCA AATCCCGTGC TTCAACATAG GTGTCCCCTG | 9840 |
| TTCCATACA GGCCTGAAAA AGACTGTGAA CTAGACCGGC GCGGAAGCGT TCAGGAGCCT | 9900 |
| GACTTTCGAT GCCTAGTTCC TCAGCTAGTT GGTCAGCAAT GGTAAAGCCC AAACCTTGA | 9960 |
| TATCCTCAAC CAGCTGGTAG GGATAATTTT CAACCACATC AAGGGTTTCT TCCTTGTAAT | 10020 |
| AGTCTTGAAT CTGAAAGGCT AGTTTGTGG GAATGCCGTA GTTGGCTAGT TTGGCCAAAA | 10080 |
| TCATCTCCGT TCCGTAGTTG AGACGGAGAG TGGAGACGAA AGCCTCGCGA TTTTGGCAG | 10140 |
| AGAGTCCTGC GATGCCCTCT AACTTTTCTG GGTGTTGCAA AATTTCTGCA ATGGTATTTT | 10200 |
| CGCCATAGGT ATCCACGATT TTCTGAGCTG TCTTGAGACC AATCCCCTTG AAATGGCTAC | 10260 |
| TTGAAAAGTA CTTGACCAAG CCCTTACTAG TTGGTTTGC GCGATCATAA CGACTGATTT | 10320 |
| GCAGTTGTTT TCCATACTTG GAGTGCTGGA CAATTTGCCC CAAAAAGTA TAGTCTTCGC | 10380 |

CCTCAATTAC ATCAGCCATG GTTCCTGTGA CAATGATTTC AAAATCATCA AAATCCTCTG 10440
CGTCCGTATC GTCGATTTCT AGGAGGAGGA TCGATAAAA ATTGCTGGGA TTTTCAAAAA 10500
TAATCCGTTT AATAGTTCCT GAAAAATAAA CTTCCATAAA ATTCCTTTGC ATGAATAGGT 10560
GAGAGTTGGG ATTGTTTTTA TTTTATACTC TTCGAAAATA TCTTCAAACC ACGTCAGCTT 10620
CCATCTGCAA CCTCAAAACA GTATTTTGAG CTGACTTCGT CAGTTCTATC CACAACCTCA 10680
AAACACTGTT TTAAGCAGCC TACGGCTAGC TTCCTAGTTT GTTCTTTGAT TTTTATTGAG 10740
TATTTGTAAA TAAACAATCA CTTCTCACGA TAGAAGAAGA GGCTGAGATT GGTGATTCTC 10800
TGCTCTTAG GTTCTTAAA ATGTTCCGAT ACGGGTGATT GGCCATAAGC GGAATTTAGC 10860
TTCCCTGTG ATATCTTTTG CTTTGAAGGT ACCTACGTGG CGGCTGTCGC TCGAAACCAA 10920
GCGGTCATCT CCGAGGAGAA GGTATTCTCC TTCTGGAACA GTAAAGCTAA AGTTGGTGT 10980
GTAGTTGACA TCAACTGTGA AGGCTTGAGC TTTTGTAGCG ATACTTCTAA AGAAAGTTCC 11040
TTTATTTCTT TCAAAGCCCT TGCCTGAGTA AGTGCTTTGG AGTTTGTCAT CCTTGAAGCG 11100
TTTGATATAG TCTGCTAGAT AAGGCTCGTC CGTTTCTTTG TCATTGATGT AGAGTTTATC 11160
ATTTTCGTAA CGAATGGTGT CGCCAGGCAT TCCAATCAGC CGCTTGACGA TGTCTTATT 11220
GCCATCTTCC TCATGGGCCA CCACGATATC AAAACGGTCA ATAGGAAGGT GTTTTACAAC 11280
GAAGAGAATT TCGCCATCCG CTAGGGTCGG ATCCATGGAA TGTCTTCTA CGCGAACATT 11340
GCTCCAAAAA AAGATACGAC TTAAAGCTAG TAATGACAGA ATTAGGAGGA ACAATCCCCA 11400
CTCTTTTAAG AAATTTTAA ATGAATTCAT AACTTACCTT TCTAAGCGTT TTTTCGCTTT 11460
TTCAGTGTTT TTAAAGTGCA ATTTGGCGCA GAAGCTGAGT CCCTGCATAC CATAGGCTTG 11520
CAAAATCTGG CTAGCCACCT TGTGAGAAGC CGTTCAGCT CCACTTGGGA GCTGATAACC 11580
CAGTTCTCGT CCCAAATTTT CAAGATTTTC CAGAAAGAGA TCACGCGCAA TGACAGAAGA 11640
AACTGCGACA GACAAGTATT TGCCCTCAGC CTTTCTTCT AAGCTGATAG GATTGCTGAA 11700
ACGATTGGCC TCTGTGCCA AGTACTGTG ATAATTTTGA GCACTGGTAA AGGCATCAAT 11760
CACAATTTTC TCAGGCTGAA CACCTTTTTG AAGGAGGAGA TAGATAGCCT GATTATGGAG 11820
GGCAACCTTA ACCGAAACAG CGTTGTAGCG GTCTCCGATG ACCTCGTTGT ACTTGCTGGG 11880
TGAGAGAAGG AGTGCCTGGT GCTGAATTTT TTCCTTGAGA ATAGGAGTAA TCTGACGGAT 11940
CTTTTGGTCG GTCAGAGTCT TAGAATCCCC CACACCGAGT TTTTCGTAAA AGTCGTGCTG 12000
GTCAGGTGTG ACAAAGGCAG CCACAACCTGC AAGCCACCA AAGTAGGAAC CATTTCCAC 12060
CTCATCTGTC CCAATTAAAG GAAGATTTTG TCCGCTGGTT TGCTCTACAG CTTGATAGCC 12120

658

| | |
|--|-------|
| AAAGAACTG GCGTATTTT CAGCCCCCTC ACCCTGAAGC AAGATTTTTC CAGAAGTATA | 12180 |
| GATAGAAACC GTTGCTTGAG GTAGTTTCAA AAAGTAGCGG ATATAGGGAT TCTTGCTAGG | 12240 |
| AGCCAGACTG GTTTGATAGT GTTCAAGAAA AGCCTGAATA TCCTTTTCGC TTGGTGTGAG | 12300 |
| TGTGATACTT GCCATAGTTT CTATTGTACC ACAAAGCAG TAAAATTTGT AAAAAGTAC | 12360 |
| AAAAATAGCG AATTTTGGA TAATATCGTG AGGTGAATTT TATGGCAAAT CTAAATCGAT | 12420 |
| TCAAATTTAC ATTCGGGAAA AAATCGTTAA CCTTGACAAG CGAACATGAC AACCTTTTTA | 12480 |
| TGGAGGAAAT CGCTAAGGTT GCGACAGAAA AATACCAAGC AATTAAAGAA CAAATGCCTA | 12540 |
| GCGCAGATGA TGAAACAATC GCTCTTTTGT TGGCAGTCAA CTGTTTATCA ACTCAGCTCA | 12600 |
| GCCGTGAGAT TGAATTTGAC GATAAGGAGC AAGAGCTAGA AGAACTCCGT CACAAGCTTG | 12660 |
| TGACTTGTA GCAAGAACAG AGCAAGATTG AGGATTCCTT ATGATTTTCAT TCCTTCTTCT | 12720 |
| ATTGGTCTTG GTTTGGGGAT TTTATATCGG CTATCGGAGA GGCCTGCTCT TACAGGTTTA | 12780 |
| TTACCTGATT TCAGCCATGG CATCGGCTTT TATGGCTGGC CAGTTTATA AGGGGCTTGG | 12840 |
| AGAGCAATTC CATTTATTGC TCCCTTATGC AAATTCGCAG GAAGGTCAGG GGACTTCTT | 12900 |
| TTTCCCATCG GATCAACTCT TTCAGCTGGA TAAGGTCTTT TATGCAGGTA TCGGCTACTT | 12960 |
| GCTTGATTT GGGATTGTCT ATAGCATTTG TCGTTTACTT GGTCTTCTCT TACACTTGAT | 13020 |
| TCCTAGCAAA AAAGTGGGTG GTAAGTTGTT CCAAGTTTCA GCAGGTATCT TGTCCATGTT | 13080 |
| GGTGACCTTA TTTGTCTTGC AAATGGCCTT GACAATCTTG GCGACCATCC CCATGGCAGT | 13140 |
| TATACAAAAT CCTCTTGAAA AGAGTATCGT CGCAAAACAC ATCATCCAGA GCATACCGGT | 13200 |
| AACAACCACT TGGCTCAAAC AAATCTGGGT GACAAATTTA ATCGGATAAA AAGGGCAGGA | 13260 |
| GTTTTCTTAG CCCTTTGTTT ACAGATTTGA CTCGAATCTA TCAGAATGTA AAAAGCTACC | 13320 |
| ACACCTAGAC ATTCAAAGAC AAGGAAATAA AGATGAATAA GAAAATATTA GAAACATTAG | 13380 |
| AGTTCGATAA GGTCAAGGCC TTGTTTGAGC CTCATTGTGTT GACCGAGCAG GGCTTGGAGC | 13440 |
| AATTGAGACA ACTGGCTCCG ACTGCCAAAG CAGATAAAAT CAAACAGGCT TTTGCTGAGA | 13500 |
| TGAAGGAAAT GCAGGCTCTT TTCGTCGAGC AACCGCATT TACTATTCTC TCAACTAAGG | 13560 |
| AAATGCGAGG AGTCTGCAAG AGGTTGGAGA TGGGAGCGGA TCTCAATATC GAGGAGTTCC | 13620 |
| TACTCTTGAA ACGCGTGCTT CTTGCCAGCC GAGAACTTCA AAATTTTAC ACCAATCTGG | 13680 |
| AAAATGTCAG CTTGGAAGAA TTAGCCCTTT GGTGAGAA ATTACATGAT TTTCCGCAAT | 13740 |
| TACAAGGAAA TCTTCAGGCC TTAAATGATG CGGGTTTCAT TGAAAATTTT GCCAGTGAAG | 13800 |
| AATTGGCGCG AATCCGTCGA AAAATACATG ATAGCGAGAG TCAGGTACGC GATGTTTTAC | 13860 |
| AAGACTTGCT CAAGCAAAA GCGCAGCTGT TGACGGAAGG AATTGTTGCT AGCAGAAATG | 13920 |

659

| | |
|--|-------|
| GCCGTCAGGT TTTACCAGTC AAAAACACCT ACCGCAATAA GATTGCAGGT GTCGTTTCATG | 13980 |
| ATATTCTGTC TAGTGGAAC ACCGTCTATA TCGAACCCCG TGAGGTAGTC AAAGTGAGCG | 14040 |
| AAGAAATTGC TAGTCTGCGA GCAGATGAGC GCTATGAAAT GCTTCGCATT CTCCAAGAAA | 14100 |
| TTTCTGAGCG TGTCCGCCCT CATGCGGCTG AGATTGCTAA TGACGCTTGG ATTATCGGTC | 14160 |
| ATCTGGACTT GATTCGTGCC AAGGTTTCGAT TTATCCAAGA AAGACAAGCA GTCGTGCCTC | 14220 |
| AGCTGTCAGA AAATCAAGAG ATTCAACTGC TCCATGTCTG CCATCCTTTG GTCAAAAATG | 14280 |
| CCGTCGCAAA TGATGTCTAT TTTGGTCAAG ATTTAACAGC TATTGTCATT ACAGGTCCCA | 14340 |
| ATACAGGTGG GAAGACCATC ATGCTCAAAA CTCTGGGCTT GACACAGGTC ATGGCCAGT | 14400 |
| CAGGATTGCC GATTTTAGCA GACAAGGAA GTCGTGTTGG TATTTTGA GAAATCTTTG | 14460 |
| CTGATATTGG AGATGAGCAG TCTATTGAGC AGAGCTTGTC TACCTTCTCT AGTCATATGA | 14520 |
| CCAATATCGT GGATATTCTT GGCAAGGTCA ACCAACATTC ACTCTTACTT TTGGATGAGT | 14580 |
| TGGGGGCTGG TACTGATCCC CAAGAGGGAG CAGCCCTGTC CATGGCTATT CTGGAGGACC | 14640 |
| TTCCGCTGCG TCAAATCAAG ACCATGGCGA CGACCCACTA TCCAGAATC AAGGCCTACG | 14700 |
| GTATTGAGAC AGCCTTTGTG CAAAATGCCA GTATGGAGTT TGATACTGCA ACTCTTCGCC | 14760 |
| CGACCTATCG CTMTATGCAG GGTGTTCTTG GCCGAAGTAA TGCCTTTGAA ATTGCCAAAC | 14820 |
| GTCTAGGCCT ATCTGAAGTT ATCGTAGGAG ATGCCAGTCA GCAGATCGAT CAGGACAATG | 14880 |
| ACGTCAATCG TATCATGAG CAATTAGAAG AGCAGACGCT GGAAAGCCGC AAACGTTTGG | 14940 |
| ACAAATATCCG TGAGGTGGAG CAAGAAAATC TCAAGATGAA CCGTGCCTA AAAAACTCT | 15000 |
| ACAACGAGCT TAATCGTGAA AAGGAAACCG AGCTTAACAA GGCGCGTGAA CAGGCTGCTG | 15060 |
| AGATTGTGGA TATGGCCCTA AGTGAAAGTG ACCAGATTCT CAAAAATCTC CACAGTAAAT | 15120 |
| CCCAACTCAA GCCCCACGAA ATCATTGAAG CCAAGGCCAA GTTGAAAAAA TTGGCTCCTG | 15180 |
| AAAAAGTGGA CTTGTCTAAA AATAAGGTCC TTCAAAAGGC CAAGAAAAAA CGAGCTCCAA | 15240 |
| AGGTGGGAGA TGATATCGTG GTTCTCAGTT ATGGTCAGCG TGGTACCTTG ACCAGTCAAC | 15300 |
| TCAAGGACGG TCGCTGGGAA GCCCAAGTTG GCTTGATTAA GATGACCTTG GAAGAGAAAG | 15360 |
| AGTTTGATCT TGTTCAGCC CAGCAAGAAA AACCAGTCAA GAAGAAACAG GTCAATGTTG | 15420 |
| TGAAACGAAC TTCTGGGCGA GGACCTCAAG CTAGACTGGA TCTTCGAGGC AAGCGCTATG | 15480 |
| AAGAAGCCAT GAATGAGCTA GATACCTTCA TCGACCAAGC CTTGCTTAAC AATATGGCTC | 15540 |
| AAGTTGATAT CATCCATGGT ATCGGAACAG GAGTCATCCG TGAAGGAGTT ACCAAATACT | 15600 |
| TGCAAAGAAA CAAACATGTC AAGAGTTTCG GCTATGCCCC ACAAATGCT GGAGGCAGTG | 15660 |

660

| | |
|---|-------|
| GTGCGACTAT TGTCACTTTT AAAGGATAGC AGTATTCTGG ACTTTATAAA GTAAAAACTG | 15720 |
| TTGAACTAAT TTTTACTAAT AAACACATTG ACAAAGCCA ACATTTTGTG TAAAATTAGA | 15780 |
| ATCAATTAAA TACCAACACC GAATGAAGTT TAATAGAAGT GGGGAATCGT TTGATTTTCC | 15840 |
| ATGACTGTAA ATGGACGGAA CTCTGGAGAG ACCGTAAAGG CACCGAAGGG CAAGGCAGGC | 15900 |
| AAC TGCTCAA ACTCTCAGGT AAAAGGACAG AGCTAGGATA GACCGCTTTT TAGCATTAT | 15960 |
| CTAAGCATTC CAGAGTACAT GTATCTTGCA TGTGCTCTTT CTTTGGGGT TGAAACGATA | 16020 |
| GGAGAAGGAA ATGTTAGAAT TGCTTAAATC AATCGATGCT TTTGCTTGGG GACCGCCCT | 16080 |
| CTTGATTTTA TTGGTCGGAA CAGGGATTTA CCTAACTATT CGGCTAGGAC TCTTGCAGGT | 16140 |
| TTTGCGTCTA CCCAAGGCCT TTCAGCTTAT TTTTATCCAG GATAAGGGAC ATGGTGATGT | 16200 |
| ATCCAGTTT GCAGCTCTGT GTACAGCCTT GGCATCAACT GTTGAACAG GAAATATCAT | 16260 |
| AGGAGTTGCG ACGGCTATCA AGGTTGGTGG ACCAGGAGCT CTATTTTGA TGTGGATGGC | 16320 |
| GGCTTTCTTT GGAATGGCTA CCAAGTATGC GGAAGGACTC TTGGCCATCA AATACCGCAC | 16380 |
| CAAGGACGAC CATGGTGCAG TAGCGGGAGG TCCCATGCAT TATATCCTTC TAGGGATGGG | 16440 |
| AGAAAAGTGG CGACCACTTG CTGTTTTGTT TGCAGTAGCA GGAGTATTGG TTGCTCTCTT | 16500 |
| GGGAATCGGA ACCTTCACCC AAGTCAACTC GATTGCAGAA TCTATCCAAA ATACAACGAC | 16560 |
| GATTTGCGCA GCCATCACAG CTCTCGTCTT GTCTGTCTTT GTAGCGATTG CAGTCTTTGG | 16620 |
| TGGACTCAAG TCTATTTCTA AGGTTTCAAC TACTGTTGTT CCTTTTATGG CCATCATTTA | 16680 |
| TATCTTAGGA ACTCTTACAG TTATTTTCTT TAATATCGGA AAAATCCCTG GCACAATCGC | 16740 |
| TTTAGTCTTT ACCTCAGCTT TTAGTCCCCT TGCTGCGGTA GGTGGATTG CTGGTGCTAG | 16800 |
| CGTTCGGATG GCTATTCAAA ATGGTGTGGC GCGTGGTGTG TTCTCAAACG AATCTGGTCT | 16860 |
| GGGTTCTGCT CCTATTGCAG CTGCAGCTGC CAAGACAAAT GAACCAGTAG AGCAAGGTTT | 16920 |
| GATTTCCATG ACAGGAACCT TTATTGATAC CCTCATCATT TGTACTCTAA CTGGTTTGAC | 16980 |
| CATCTTGGTA ACTGG | 16995 |

(2) INFORMATION FOR SEQ ID NO: 83:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 28473 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 83:

| | |
|---|----|
| CCGGGGCTTT TGTAGTATAA TAGAGATACG TTTTGAAAGT AGGAGGTATC TATGGACTTA | 60 |
|---|----|

661

| | |
|--|------|
| ACTAAGCGCT TTAATAAACA GTTAGATAAA ATTCAAGTTT CGTTGATTCG TCAGTTTGAC | 120 |
| CAGGCTATTT CGGAGATTCC TGGGGTCTTG CGTTTGACCT TGGGGGAACC TGATTTTACA | 180 |
| ACGCCAGACC ATGTCAAGGA GGC GGCAAG CGAGCGATTG ATCAGAACCA ATCCTACTAT | 240 |
| ACAGGGATGA GTGGTCTGCT GACTCTACGT CAGGCAGCCA GTGACTTTGT TAAGGAAAAG | 300 |
| TACCAACTGG ACTATGCTCC TGAAAATGAA ATCTTGTTA CAATTGGGGC GACAGAGGCT | 360 |
| TTATCTGCGA CTTTGACGGC TATTTTGAA GAGGGAGACA AGGTACTTTT GCCAGCTCCT | 420 |
| GCTTATCCAG GCTATGAACC GATTGTTAAC TTAGTTGGGG CAGAAATTGT TGAGATTGAT | 480 |
| ACGACTGAAA ATGGTTTTGT CTTGACTCCT GAGATGTTGG AGAAGGCCAT TTTGGAGCAG | 540 |
| GGTGATAAGC TCAAGGCGGT TATCTCAAC TATCCAGCCA ATCCGACAGG AATTACCTAC | 600 |
| AGTCGAGAGC AGTTAGAGGC CTTGGCAGCT GTTTTACGCA AGTACGAAAT TTTGTTGTC | 660 |
| TGTGATGAGG TTTACTCAGA ATTGACCTAC ACAGGCGAAG CCATGTGTCT CTAGGAACGA | 720 |
| TGTTGAGAGA CCAGGCTATT ATTATCAATG GTTTGTCTAA ATCGCATGCC ATGACAGGTT | 780 |
| GGCGTTTGGG GCTGATTTTC GCTCCTGCGA CCTTCACAGC CCAGTTAATC AAGAGTCACC | 840 |
| AGTACTTGGT CACTGCCGCA AATACCATGG CGCAACATGC TGCGGTAGAA GCCTTGACGG | 900 |
| CTGGTAAAAA CGATGCGGAC CCATGAAGAA GGAATATATC CAACGTCGGG ACTATATCAT | 960 |
| CGAAAAATG ACTGCTCTTG GTTTTGAGAT TATCAAACCA GACGGTGCCT TCTATATTTT | 1020 |
| TGCTAAATT CCAGCGGGCT ACAATCAAGA CTCCTTTGCT TTTCTGAAGG ATTTTGCTCA | 1080 |
| GAAGAAGGCC GTTGCCCTTA TCCCTGGTGC AGCCTTTGGA CGTTACGGGG AAGGCTACGT | 1140 |
| CCGCCTATCT TATGCAGCCA GCATGGAGAC TATCAAAGAA GCCATGAAAC GACTTGAGGA | 1200 |
| GTACATGAGA GAAGCATGAT TCAGTCTATC ACGAGTCAAG GCTTGGTGCT TTACAATCGC | 1260 |
| AATTTTCGTG AGGATGACAA GCTCGTCAA ATTTTACAG AGCAGGTTGG CAAACGCATG | 1320 |
| TTTTTTGTCA AACACGCTGG TCAGTCTAAG CTGGCGCCTG TTATTCAGCC CTTGGTGCTG | 1380 |
| GCACGATTTT TCTTGCGAAT CAATGATGAC GGAATCAGTT ACATCGAAGA CTATCATGAG | 1440 |
| GTCATGACTT TCCCCAAGAT TAATAGTGAC CTCCTTTGTCA TGGCCTATGC GACCTATGTG | 1500 |
| GCAGCTCTTG CAGATGCTAG TTTGCAGGAC AATCAGCAGG ATGCTCCCTT GTTTGCTTTT | 1560 |
| TTGCAAAAGA CTTTGAGATT GATGGAAGCA GGCTTGGATT ATCAGGTTTT GACCAATATT | 1620 |
| TTTGAAATTC AAATTTTGAC TCGATTTGGA ATCAGCCTCA ATTTTAATGA GTGTGTCTTC | 1680 |
| TGCCATCGGG TTGGTCAGGC TTTTGACTTT TCTTTCAAAT ATGGAGCCTG CCTCTGTCCA | 1740 |
| GAGCATTATC ATGAGGATAA GAGACGTTGT CATCTCAATC CCAATATCCC CTATCTGCTC | 1800 |

| | | | |
|------------|-------------|-------------|-------------|
| | | 662 | |
| AATCAATTTC | AAGCTATTGA | TTTGTGAGACT | TTGGAGACCA |
| 1860 | TTTCGCTCAA | GCCTGGAATC | |
| AAGCAAGAGC | TACGCCAATT | TATGGATCAA | TTATATGAAG |
| 1920 | AGTACGTTGG | GATTACACCTA | |
| AAATCAAAGA | AATTTATTGA | TTCCCTAGCA | GACTGGGGAC |
| 1980 | AATTACTAAA | AGAGGAAAAG | |
| AAATGAAAAA | AATCGCAGTA | GATGCCATGG | GGGGCGATTA |
| 2040 | CGCACCTCAG | GCCATTGTTG | |
| AGGGTGTCAA | TCAAGCCCTA | TCTGACTTTT | CAGATATCGA |
| 2100 | GGTTCAACTT | TACGGAGATG | |
| AAGCTAAAAT | CAAGCAATAT | CTGACAGCGA | CAGAGCGCGT |
| 2160 | CAGCATTATC | CATACGGATG | |
| AGAAGATTGA | TTCGGATGAT | GAACCTACGA | GAGCTATTCG |
| 2220 | GAATAAGAAA | AATGCCAGTA | |
| TGGTATTGGC | AGCCAAGGCT | GTCAAAGATG | GTGAAGCAGA |
| 2280 | CGCTGTCCTT | TCGGCTGGGA | |
| ATACAGGTGC | CTTGTGGCA | GCAGGATTCT | TCATCGTGGG |
| 2340 | TCGTATCAAG | AATATCGACC | |
| GTCTTGGACT | CATGTCTACC | TTGCCTACCG | TTGATGGAAA |
| 2400 | AGGTTTGTGAC | ATGCTAGACC | |
| TTGGTGCCAA | TGCAGAAAAT | ACAGCCCAGC | ACCTCCATCA |
| 2460 | ATATGCGGTT | CTAGGTTCTT | |
| TCTATGCTAA | AAATGTCCGT | GGCATTGCGC | AACCACGCGT |
| 2520 | TGGTTTGCTC | AACAACGGAA | |
| CAGAGAGTAG | CAAGGGCGAC | CCGCTTCGTA | AGGAACTTA |
| 2580 | TGAATTACTG | GCGGCTGATG | |
| AAAGTTTGAA | CTTTATCGGA | AACGTGGAAG | CGCGTGATTT |
| 2640 | GATGAATGGC | GTTGCAGATG | |
| TTGTTGTGGC | AGATGGTTTC | ACGGGAAACG | CTGTGCTCAA |
| 2700 | ATCCATCGAA | GGGACAGCTA | |
| TGGGAATCAT | GGGCTTGCTC | AAGACAGCTA | TTACAGGTGG |
| 2760 | TGCTCTTCGA | GCGAACTAG | |
| GTGCCCTCCT | TCTCAAGGAC | AGCCTCAGTG | GTTTGAAAAA |
| 2820 | ACAGCTCAAT | TATTCAGATG | |
| TTGGTGAGC | GGTCTGTGTT | GGTGTTAAGG | CACCTGTTGT |
| 2880 | CAAGACTCAT | GGCTCAAGCG | |
| ATGCCAAGGC | TGTTTATAGT | ACGATTTCGT | AGATCCGTAC |
| 2940 | CATGCTAGAA | ACAGACGTGG | |
| TTGCCCAGAC | TGCGCGTGAA | TTTTCAGGAG | AATAAAAGAG |
| 3000 | ATGACAGAAA | AAGAAATTTT | |
| TGACCGTATT | GTGACCATTA | TCCAAGAGCG | ACAGGGAGAG |
| 3060 | GACTTTGTCT | TGACAGAATC | |
| CTTGAGTCTG | AAAGACGATT | TGGATGCCGA | TTCTGTTGAC |
| 3120 | TTGATGGAGT | TTATCTTGAC | |
| TCTGGAAGAT | GAATTTAGTA | TCGAAATCAG | CGATGAAGAA |
| 3180 | ATTGACCAAC | TCCAAAACG | |
| AGGAGATGTG | GTTAAAATCA | TTCAAGGAAA | ATAGCAATCG |
| 3240 | GAGTTCCAAG | TCAACGGAAG | |
| TAGATGGTTT | TTAGAAATGA | GAAATATCGG | ACAAGCTGGT |
| 3300 | AAAATCTTGG | CTGACAGTGG | |
| TTATCAAGGG | CTCATGAAGA | TATATCCTCA | AGCACAAACT |
| 3360 | CCACGTAAAT | CCAGCAAAC | |
| CAAGCCGCTA | ACAGTTGAAG | ATAAAGCCTG | TAATCATGCG |
| 3420 | CTATCTAAGG | AGATAAGCAA | |
| GGTTGAGAAT | ATCTTTGCCA | AAGTAAAAAC | GTTTAAAAATG |
| 3480 | TTTTCAACAA | CCTATCGAAA | |
| TCATCGTAAA | CGCTTCGGAT | TACGAATGAA | TTTGATTGCT |
| 3540 | GGTATTATCA | ATCATGAACT | |
| AGGATTCTAG | TTTTCAGGA | AGTCTAATAG | TAAAAAAGTG |
| 3600 | ATTAGAAAAC | ATCTTTTCTTA | |

663

| | |
|--|------|
| AAAATAGAGA TGATTTTGAA ACAAAAAAGC TAATTCAAGA CGTTTCGATG CCAATTCAAG | 3660 |
| ATTTGGATGA AAAAAATTAA TAGATACTGT TATACTAAAC TTGTCAAGTT TGTAACAAGA | 3720 |
| CAAATATTAA AAATAAAAAA GAGGTATTCTG TTATGAATAC AAAAACGATG TCACAATTG | 3780 |
| AAATTATGGA TACTGAGATG CTTGCTTGCG TTGAAGGTGG CGGATGCAAT TGGGGAGATT | 3840 |
| TTGCCAAAGC AGGTGTTGGA GGAGGAGCAG CACGAGGTCT TCAGCTAGGA ATTAAAAACA | 3900 |
| GAACATGGCA AGGTGCAGCA ACTGGTGCTG TGGGAGGAGC TATACTTGGA GGTGTGGCCT | 3960 |
| ATGCAGCGAC ATGTTGGTGG TAATTATGGA TTTTAAAAGT TTTATTATTG GTTTAGTAGT | 4020 |
| TGGTATATTT GGTCTTATA TGGATGATTT AATTAGAAAA AAATTTTAA AGTCTTCGGA | 4080 |
| GAAGAAAACA GAAAAATCTG TTAAAAATA ATCAAACTA TAAATGATGA ATCTGAATCA | 4140 |
| AAATTATTTT GCGCATGTAA AGAGGAGTCT TATAGTAACG AGTCAAAAAA GGAGTAACTA | 4200 |
| TGAATCGTAA TTTAGAACGG TGTATCTAT TCTGACTAGG AATAGATCAT ACCAGAGGTA | 4260 |
| GCTTAGAAAT AGCAGAGACA TTAGAAATTG AAGTAATAAA TAGGATGTCG TAAGTGTTAC | 4320 |
| TATCAATGAT TTATTTGTTT CAAGCTTGCC TAGGGTGACA GTAAAAATC AATTCCTTT | 4380 |
| CAATAGCATA TTTTATAGTG GCAGGACTCT TGTCTGCCT ATTTTMTTAT CCAAAAAGTG | 4440 |
| CAGTTGGGAG GGAGATAGGC TCATTGTTGGA AGGAAGTCCA GTTTTGTGTT AGTGATTGGG | 4500 |
| GTAAGATAGT TGTATCAGA TGAGTTAATA CTCTCGAAA ATCAAATTCA AACCACGTCA | 4560 |
| ACGTCGCCTT GCCGTATATA TGTGACTGAC TTCGTCAGTC CTATCTACAA CCTCAAAACA | 4620 |
| GTGTTTGTAG CAGCCTACGG CTAGTTTCTT AGTTTGCTCT TTGATTTTCA TTGAGTATTA | 4680 |
| GGGAAAAGGA GATGAATATG AAATTTGGGA AACGTCATTA TCGTCCGCAG GTGGATCAGA | 4740 |
| TGGACTGCGG TGTAGCTTCA TTAGCCATGG TTTTGGCTA CTATGGTAGT TATTATTTT | 4800 |
| TGGCTCACTT GCGAGAATTG GCTAAGACGA CCATGGATGG GACGACGGCT TTGGGCTTGG | 4860 |
| TCAAGGTGGC AGAGGAGATT GGTTTTGAGA CGCGAGCCAT TAAGGCAGAT ATGACGCTTT | 4920 |
| TTGACTTGCC GGATTTAACT TTTCTTTTG TTGCCCATGT GCTTAAGGAA GGGAAATTGC | 4980 |
| TCCACTACTA TGTGGTGACT GGGCAGGATA AGGATAGCAT TCATATTGCC GATCCAGATC | 5040 |
| CCGGGGTGAA GTTGACTAAA CTGCCACGTG AGCGTTTGA GGAAGAATGG ACAGGAGTGA | 5100 |
| CTCTTTTAT GGCACCTAGT CCAGACTATA AGCCTCATAA GGAACAAAAA AATGGTCTGC | 5160 |
| TCTCTTTAT CCTATATTA GTGAAGCAGC GTGGCTTGAT TGCCAATATC GTTTGGCAA | 5220 |
| CACTCTTGGT AACCGTGATT AACATTGTGG GTTCTTATTA TCTGCAGTCT ATCATTGATA | 5280 |
| CCTATGTGCC AGATCAGATG CGTTCGACAC TAGGGATTAT TTCTATTGGG CTAGTCATCG | 5340 |

664
TCTACATCTT CCAGCAAATC TTGTCTTACG CTCAGGAGTA TCTCTTGCTT GTTTTGGGGC 5400
AACGCTTGTC GATTGACGTG ATTTTGTCTT ATATCAAGCA TGTMTTTCAC CTCCCTATGT 5460
CCTTCTTTGC GACACGCAGG ACAGGGGAGA TCGTGTCTCG TTTTACAGAT GCTAACAGTA 5520
TCATCGATGC GCTGGCTTCG ACCATCCTTT CGATTTTCCT AGATGTGTCA ACGGTTGTCA 5580
TTATTTCCCT TGTTCATTT TCACAAAATA CCAATCTCTT TTTTCATGACT TTATTGGCGC 5640
TTCTTATCTA CACAGTGATT ATCTTTGCCT TTATGAAGCC GTTTGAAAAG ATGAATCGGG 5700
ATACCATGGA AGCCAATGCG GTTCTGTCTT CTCTATCAT TGAGGACATC AACGGTATTG 5760
AGACTATCAA GTCCTTGACC AGTGAAGTC AGCGTTACCA AAAAATTGAC AAGGAATTTG 5820
TGGATTATCT GAAGAAATCC TTTACCTATA GTCGAGCAGA GAGTCAGCAA AAGGCTCTGA 5880
AAAAGGTTGC CCATCTCTTG CTTAATGTG GCATTCTCTG GATGGGGGCT GTTCTGGTCA 5940
TGGATGGCAA GATGAGTTTG GGGCAGTTGA TTACCTATAA TACCTTGCTG GTTTACTTTA 6000
CTAATCCTTT GGAAATATC ATCAATCTGC AAACCAAGCT TCAGACAGCG CAGGTTGCCA 6060
ATAACCGTCT AAATGAAGTG TATCTAGTAG CTCTGAGTT TGAGGAGAAG AAAACAGTTG 6120
AGGATTTGAG CTTGATGAAG GGAGATATGA CCTTCAAGCA GGTTCAATTAC AAGTATGGCT 6180
ATGGTCGAGA TGTCTTATCG GATATCAATT TAACCGTTCC CCAAGGGTCT AAGGTGGCTT 6240
TTGTGGGGAT TTCAGGGTCA GGTAAGACGA CTTTGGCCAA GATGATGGTT AATTTTACG 6300
ACCCAAGTCA AGGGGAGATT AGTCTGGGTA GTGTCAATCT CAATCAGATT GATAAAAAAG 6360
CCCTGCGCCA GTACATCAAC TATCTGTCTC AACAGCCCTA TGTCTTTAAC GGAACGATT 6420
TGGAGAATCT TCTTTTGGGA GCCAAGGAGG GGACGACACA GGAAGATATC TTACGGGCGG 6480
TCGAATTGGC AGAGATTCTG GAGGATATCG AGCGCATGCC ACTGAATTAC CAGACAGAAT 6540
TGACTTCGGA TGGGGCAGGG ATTTACAGTG GTCAACGTCA GAGAATCGCT TTGGCGCGTG 6600
CTCTCTTGAC AGATGCGCCG GTCTTGATTT TGGATGAGG GACTAGCAGT TTGGATATTT 6660
TGACAGAGAA GCGGATGTG GATAATCTCA TTGCTTTGGA CAAGACCTTG ATTTTCATTG 6720
CTCACCCTT GACTATTGCT GAGCGGACAG AGAAGGTAGT TGTCTTGGAT CAGGGCAAGA 6780
TTGTGGAAGA AGGAAAGCAT GCTGATTTGC TTGCACAGGG TGGCTTTTAC GCCCATTTGG 6840
TCAATAGCTA GAAAGAGGAG AGGATGAAAC CAGAATTTT AGAAAGTGCG GAGTTTATA 6900
ATCGTCGTTA CCATAATTT TCCAGTAGTG TGATTGTACC CATGGCCCTT CTGCTTGTGT 6960
TTTACTTGG CTTTGCACT GTTGCAGAGA AGGAGATGAG TTTGTCCACT AGAGCTACTG 7020
TCGAACCTAG TCGTATCCTT GCAAATATCC AGTCAACTAG CAACAATCGT ATTCTTGTCA 7080
ATCATTTGGA AGAAAATAAG CTGGTTAAGA AGGGGATCT TTTGGTTCAA TACCAAGAAG 7140

665

| | |
|--|------|
| GGGCAGAGGG TGTCCAAGCG GAGTCCTATG CCAGTCAGTT GGACATGCTA AAGGATCAAA | 7200 |
| AAAAGCAATT GGAGTATCTG CAAAAGAGCC TGCAAGAAGG GGAGAACCAC TTTCCAGAGG | 7260 |
| AGGATAAGTT TGGCTACCAA GCCACCTTTC GCGACTACAT CAGTCAAGCA GGCAGTCTTA | 7320 |
| GGGCTAGTAC ATCGCAACAA AATGAGACCA TCGCGTCCCA GAATGCAGCA GCTAGCCAAA | 7380 |
| CCCAAGCCGA AATCGGCAAC CTCATCAGTC AAACAGAGGC TAAAAATTCGC GATTACCAGA | 7440 |
| CAGCTAAGTC AGCTATTGAA ACAGGTGCTT CCTTGGCCGG TCAGAATCTA GCCTACTCTC | 7500 |
| TTTACCAGTC CTACAAGTCT CAGGGCGAGG AAAATCCCCA AACTAAGGTT CAGGCAGTTG | 7560 |
| CACAGGTGA AGCACAGATT TCTCAGTTAG AATCTAGTCT TGCTACTTAC CGTGTCCAGT | 7620 |
| ATGCAGGTTT AGGTACCCAG CAAGCCTATG CGTCAGGGTT AAGCAGTCAA TTGGAATCCC | 7680 |
| TTAAATCCCA ACACTTGGCA AAGGTTGGTC AGGAATTGAC CCTTCTAGCC CAGAAAATTT | 7740 |
| TGGAGGCAGA GTCAGGTAAG AAGGTACAGG GAAATCTTTT AGACAAGGGG AAAGTTACGG | 7800 |
| CGAGTGAGGA TGGGGTGCTT CATCTTAATC CTGAGACCAG TGATTCTAGC ATGTTGCAG | 7860 |
| AAGGTGCCCT ACTAGCCCAA CTTTATCCAT CTTTGGAAAG AGAAGGGAAA GCCAACTCA | 7920 |
| CAGCTTATCT AAGTTCAAAA TATGTAGCAA GAATCAAGGT CGGTGATTCT GTTCGCTATA | 7980 |
| CTACGACTCA TGATGCCGGG AATCAACTTT TCCTAGATTG TACTATTACA AGTATTGATG | 8040 |
| CGACAGCTAC TAAGACTGAG AAAGGGAATT TCTTTAAAT CGAGGCGGAG ACTAATCTAA | 8100 |
| CTTCGGAGCA GGCTGAAAAA CTTAGGTACG GGGTGAAGG CCGCTGCAG ATGATTACGG | 8160 |
| GCAAGAAAAG TTACCTACGT TATTATTTGG ATCAATTTTT GAACAAAGAG TAATGTTCGT | 8220 |
| GTTTTTAGAG TTAAATAATT TTTAACTGT GAGAAAGATT CTTCTTGCAG TTTTTTCTTT | 8280 |
| ACAATTTTTG AAAACATCT ACTATTTATT CGTTAAAT CTTGTGTTTT TTGGTTTTTT | 8340 |
| GTGGTAAAT GTGCTCAAGT AATACGAAAG GCGAACTTTA AAATGTCAA ACAATTGATC | 8400 |
| TATTCGGGAA AAGCTAAAGA TATCTATACA ACTGAGGATG AAAATCTTAT TATTTCAACT | 8460 |
| TACAAGGACC AGGCGACTGC TTTCAACGGT GTCAAGAAGG AGCAGATTGC AGGTAAGGGA | 8520 |
| GTCTTGAATA ATCAGATCTC ATCTTTTATT TTTGAGAAAT TAAATGTGGC TGGTGTGGCG | 8580 |
| ACTCACTTTG TGGAGAACT TTCAGACAGG GAACAACTCA ATAAAAAGGT TAAGATTATT | 8640 |
| CCTTTGGAAG TCGTGCTCCG CAACTATACT GCTGGTTCCT TTTCAAAACG TTTTGGTGTG | 8700 |
| GATGAGGGAA TCGCCTTGGA GACTCCGATT GTCGAATTTT ACTACAAAA TGATGATTG | 8760 |
| GATGATCCAT TTATCAATGA TGAGCATGTG AAATTCCTAC AGATTGCGGG TGACCAGCAG | 8820 |
| ATTGCCTACT TGAAGGAAGA AACGCGTCGT ATCAATGAAC TATTGAAAGT CTGGTTTGCT | 8880 |

666

| | |
|--|-------|
| GAGATTGGGC TTAAATTGAT TGACTTTAAG CTAGAGTTCG GTTTTGACAA GGATGGCAAG | 8940 |
| ATTATCTTGG CAGACGAATT TTCACCAGAT AACTGCCGCT TGTGGGACGC TGATGGCAAC | 9000 |
| CACATGGATA AGGATGTTTT CCGTAGAGGA TTGGGAGAAC TAACCGACGT TTATGAGATT | 9060 |
| GT TTGGGAAA AGTTGCAGGA ATTGAAATAA TCTGTTTGCA ACGGAAAACC TTCGTCTCTC | 9120 |
| AACTAAAAGG ACTCAGGCTG AAAAGGTCCC CCAGACCTTT TCACTCTGTA GAGAAGTAGG | 9180 |
| TGAACTAACA GATGTTTACG AAATTGTCTG GGAAAAGTTG CAGGGTTTAA AATAACAACC | 9240 |
| TCAAGGCTGT TTGGGAATAT TGCAAGAGCT GAAATAAAGG AATAAGAATT GATGGATAAA | 9300 |
| CGTATTTTTG TTGAAAAAAA GGCTGATTTT CAGGTCAAGT CAGAGAGTTT GGT TAGAGAG | 9360 |
| CTCCAGCACA ACTTGGGACT GTCAAGCTTG AAAAGTATTC GTATTGTGCA AGTATATGAT | 9420 |
| GTATTTGACT TGGCTGAGGA CTGTTTGCA CCTGCAGAGA AGCACATTTT CTCTGAGCAG | 9480 |
| GTAACCGACC ATGTTT TAGA TGAAGTATCT GTGCAGGCGG ATCTTGCTAA CTATGCTTTC | 9540 |
| TTTGCCATTG AAAGTCTGCC AGGGCAGTTT GACCAGCGTG CAGCTTCGTC ACAGGAAGCC | 9600 |
| TTGCTTTTGT TGGGAAGTTC GAGTGACGTG ACAGTCAACA CAGCCCAACT TTA CTGGTG | 9660 |
| AATAAAGATA TTGATGCGAC TGAGTTGGAA GCTGTCAAAA ACTACCTGCT CAATCCAGTT | 9720 |
| GATTCTCGTT TCAAGGATAT CACGACAGGG ATTGCCAAGC AGGAGTTTTC AGAGTCAGAC | 9780 |
| AAGACCATT CCAAATTGAC TTTCTTTGAA AGCTATGCAG CAGAAGACTT TGCTCGCTAC | 9840 |
| AAGGCCGAAC AAGGGATGCG CATGGAAGTG GATGATTTGC TCTTTATCCA AGACTACTTT | 9900 |
| AAGTCAATCG GCGCGTGCC AACTGAGACT GAACTCAAGG TTTTGACAC TTA CTGGTCT | 9960 |
| GACCACTGCC GTCATACGAC TTTTGAGACA GAGTTGAAAC ACATCGACTT TTCAGCTTCT | 10020 |
| AAATTTCAAA AGCAATTGCA GTCAACCTAT GACAAGTATA TTGCCATGCG CGAGGAATTA | 10080 |
| GGTCGGTCTG AAAAACCACA AACCTTGATG GATATGGCGA CTATTTTCGG TCGTTATGAG | 10140 |
| CGTGCTAATG GACGATTGGA TGATATGGAA GTCTCTGACG AAATCAATGC CTGCTCAGTT | 10200 |
| GAAATTGAAG TGGACGTTGA TGGTGTCAAG GAACCTTGGC TCCTCATGTT TAAAAACGAA | 10260 |
| ACCCACAACC ATCCAACAGA AATTGAGCCA TTTGGTGGAG CGGCTACCTG TATTGGTGG | 10320 |
| GCTATTCGTG ATCCGTTGTC AGGCCGTTCC TATGTTTACC AAGCCATGCG TATTT CAGGT | 10380 |
| GCTGGTGATA TTACAGCACC GATTTTCGAA ACTCGCGCTG GGAAATTGCC ACAACAAGTC | 10440 |
| ATTTCTAAAA CAGCAGCTCA TGTTTATCT TCATATGGTA ACCAGATTGG GCTTGCAACA | 10500 |
| ACCTACGTTT GTGAATACTT CCACCCAGGC TTTGTAGCTA AACGTATGGA ACTTGGTGCC | 10560 |
| GTGTTGGTG CGACTCCCAA GGGCAATGTT GTCCGTGAAA AACCTGAAGC AGGTGATGTG | 10620 |
| ATCATCCTTC TCGGAGGCAA AACAGGTCGT GATGGTGTG GTGGTGCGAC GGGCTCTTCT | 10680 |

667

| | |
|--|-------|
| AAGGTTCAAA CAGTTGAGTC TGTAGAGACT GCTGGTGCTG AGGTTCAAAA AGGAAATGCC | 10740 |
| ATCGAAGAAC GCAAGATTCA GCGCCTCTTC CGTAATGGCA ATGTCACTCG TCTGATCAAG | 10800 |
| AAGTCCAATG ACTTTGGGGC AGGCGGTGTC TGTGTGGCTA TCGGTGAATT GGCAGACGGT | 10860 |
| CTTGAAATCG ACCTCAACAA GGTGCCTCTT AAATACCAGG GCTTGAATGG TACAGAAATT | 10920 |
| GCCATCTCTG AATCACAAGA ACGGATGGCG GTCGTGGTTC GTCCTGAAGA TGTGGATGCC | 10980 |
| TTCTGTGCCG AATGTAACAA AGAAAAATAT GATGCTGTTG TGGTGGCGAC AGTAACTGAA | 11040 |
| AAACCAAATC TTGTCATGCA CTGGAATGGT GAGACAATCG TTGACTTGGA GCGTCGTTTC | 11100 |
| CTTGACACCA ATGGTGTGCG CGTGGTTGTC GATGCCAAAG TTGTGGACAA GGATGTCAAA | 11160 |
| CTCCCAGAAG AGCGTCAAAC ATCTGCTGAA ACACTGGAAT CAGATACCCT TACGGTTCTA | 11220 |
| TCTGACCTCA ACCATGCAAG TCAAAAAGGA TTACAGACTA TCTTTGACTG CTCTGTTGGA | 11280 |
| CGCTCAACGG TTAATCACCC ACTTGGTGGT CGTTACCAAC TCACACCAAC TGAGGCATCT | 11340 |
| GTGCAGAAAT TGCCAGTTCA ACACGGTGTG ACTCATACTG CGTCGGTCAT TGCTCAAGGT | 11400 |
| TTCAACCCAT ATGTAGCTGA ATGGTCTCCA TACCACGGTG CTGCTTATGC GGTTCATCGAA | 11460 |
| GCAACTGCTC GTTTGGTGGC TGCTGGTGCC AACTGGTTCA AGGCTCGTTT CTCTTACCAA | 11520 |
| GAGTATTTTC AGCGTATGGA TAAACAAGCA GAGCGTTTCG GTCAGCCAGT AGCTGCTCTT | 11580 |
| CTAGGTTCTA TTGAAGCACA AATTGAGCTT GGCTTGCCAT CTATCGGTGG TAAGGACTCC | 11640 |
| ATGTCTGGTA CCTTTGAAGA ATTGACCGTT CCGCCAACCT TGGCTGCCTT TGGGGTGACG | 11700 |
| ACGGCAGATA GCCGTAAGGT GCTCTCTCCA GAATTTAAAG CTGTTGGGGA AAATATCTAC | 11760 |
| TACATCCCAG GTCAAGCCCT CTCTGCAGAG ATTGATTTTG ACTTGATTAA GAAAAATTTT | 11820 |
| GCTCAGTTTG AAGCCATCCA AGCTGACCAT AAAGTGACAT CTGCATCAGC TGTCAAATAC | 11880 |
| GGTGGTGTAG TTGAAAGTTT GGCTCTTGCT ACCTTTGGAA ACTATATTGG TGCAGAGGTG | 11940 |
| ACCTTGCTTG AACTTGAAAC AGCTTTGACA GCTCAATTAG GCGGCTTTGT CTTACATCT | 12000 |
| CCTGAAGAAA TTGCTGGAGT AGAGAAGGTT GGACAAACGA AAGCAGACTT TACACTGACT | 12060 |
| GTCAACGGTG TGAAGCTAGA TGGACACAAG CTTGACAGTG CATTTCAAGG GACATTGGAA | 12120 |
| GAAGTTTACC CAACAGAATT TACCCAAGCG AAAGAAGTAG AAGAAGTACC AGCTGTGGCA | 12180 |
| TCAGATGTTG TGATTAAAGC CAAAGAAAAG GTTGAAAAAC CTGTGGTTTA CATCCCAGTC | 12240 |
| TTTCCAGGAA CCAACTCAGA ATATGATTCA GCTAAGGCCT TCGAAAAAGA AGGTGCAGAG | 12300 |
| GTCAATTTGG TGCCATTTCG GACCTTGAAT GAAGAAGCTA TTGTCAAGTC AGTTGAAACT | 12360 |
| ATGGTTGACA ATATCGACAA GACTAATATT CTCTTCTTTG CTGGTGGATT CTCGGCTGCG | 12420 |

668

| | |
|--|-------|
| GATGAACCAG ATGGTTCAGC TAAGTTTATC GTCAATATCC TGCTTAATGA AAAAGTGCCT | 12480 |
| GTGGCTATTG ATAGCTTTAT CGCCCGTGGT GGTTTGATTA TCGGTATTTG TAATGGATTTC | 12540 |
| CAAGCCTTAG TCAAATCGGG TCTCCTACCC TACGGAAACT TTGAAGCTGC TAACAGTACT | 12600 |
| AGCCCAACCC TCTTCTACAA TGATGCCAAC CAACACGTGG CCAAGATGGT GGAAACTCGC | 12660 |
| ATTGCCAATA CCAACTCACC ATGGTTGGTT GGTGTGCAAG TGGGCGATAT CCACGCTATT | 12720 |
| CCTGTTTCGC ACGGTGAAGG GAAGTTTGTG GTGACGGCTG AGGAATTTGC AGAGCTCCGT | 12780 |
| GACAAATGGAC AAATTTTCAG CCAATACGTT GACTTTAACG GTAAACCAAG TATGGATTCT | 12840 |
| AAGTACAATC CGAATGGTTC TGTCATGCC ATCGAAGGAA TTACCAGCAA GAATGGTCAA | 12900 |
| ATCATCGGTA AGATGGGCCA CTCAGAACGT TATGAGGATG GTCTTTTCCA AAATATCCCA | 12960 |
| GGCAATAAAG ACCAACACCT GTTCGCATCA GCGGTTAAAC ATTTCACTGG AAAATAAGAC | 13020 |
| TTACAGATTT TCTAATAGAT AGTATCAGTA ATGTAAAAGT CATGTAAATC TAGCTCTTGA | 13080 |
| TGATTACAAA TGAAATTAG GTATAAAAA TGACATACGA AGTAAATCT CTTAATGAAG | 13140 |
| AATGTGGTGT TTTCGGTATT TGGGGACATC CAGATGCTGC TAAGTTGACC TATTTTGGAC | 13200 |
| TCCACAGTCT TCAACACCGT GGTCAAGAGG GGGCAGGAAT CCTCTCCAAT GATCAAGGAC | 13260 |
| AACTGAAGCG CCATCGTGAC ATGGGGCTTT TATCAGAAGT TTTCAGAAAT CCAGCTAATT | 13320 |
| TGGATAAATT GACAGGAGCT GGTGCGATTG GGCATGTGCG TTATGCGACT GCTGGCGAAG | 13380 |
| CTTCTGTAGA TAACATCCAG CCCTTCCTCT TCCGTTTCA CGATATGCAG TTTGGTTTGG | 13440 |
| CTCATAATGG AAATCTGACC AATGCAGCCT CTCTCAAGAA AGAACTGGAA CAAAGAGGAG | 13500 |
| CAATTTTCAG CGCGACTTCG GACTCTGAAA TCTTGGCTCA CCTCATTCGT CGCAGTCATA | 13560 |
| ATCCTAGCCT GATGGGCAA ATCAAGGAAG CGCTCAGCCT TGTCAAAGGT GGTTTTGCCT | 13620 |
| ATATCTTGCT GTTTGAGGAC AAGTTGATTG CGGCTCTTGA CCCAAATGGA TTCCGACCGC | 13680 |
| TTTCGATTGG TAAAATGGCT AATGAGCAG TTGTTGTATC TTCTGAAACC TGTGCTTTTG | 13740 |
| AGGTCAATTG TGCCGAGTGG ATTCGTGATT TGAAGCCAGG TGAGATTGTG ATCATTGATG | 13800 |
| ACGAGGGCAT TCAGTATGAC AGCTATACAG ATGATACCCA GTTGGCGGTT TGTTCATATG | 13860 |
| AGTATATCTA CTTTGCTCGC CCTGATTCTA ATATCCACGG TGTCAATGTC CATACGGCAC | 13920 |
| GTAAGAGAAT GGGAGCGCAA TTGGCGCGAG AATTTAAGCA TGAGGCAGAT ATTGTAGTTG | 13980 |
| GTGTGCCCAA TTCTTCCTA AGCGCGGCTA TGGGATTTGC GGAAGAATCA GGCTTACCAA | 14040 |
| ATGAAATGGG TCTGATCAAA AACCAATACA CCCAGCGAAC TTTTATCCAA CCGACTCAAG | 14100 |
| AATTGCGGGA GCAAGGAGTG CGGATGAAAC TGTCTGCTGT TTCGGGTGTT GTCAAAGGCA | 14160 |
| AACGTGTGGT CATGGTGGAT GATTCCATTG TACGTGGAAC AACCTCTCGT CGTATCGTTC | 14220 |

669

| | |
|--|-------|
| AGCTCTTGAA AGAAGCGGGT GCGACTGAGG TTCACGTTGC CATTGGAAGT CCTGCACTAG | 14280 |
| CGTATCCATG TTTCTACGGG ATTGATATCC AGACCCGTCA GGAGCTGATT GCAGCCAATC | 14340 |
| ATACGGTCGA AGAACTCGC CAAATCATTG GTGCGGACAG TCTGACTTAT CTTTCAATTG | 14400 |
| ATGGCTTGAT TGAGTCGATT GGTATCGAAA CAGATGCGCC GAACGGTGGT CTCTGTGTCTG | 14460 |
| CTTACTTTGA CGGTGACTAC CCAACGCCTC TTTATGACTA CGAAGAAGAC TATCGTAGAA | 14520 |
| GTTTGGAAGA AAAGACCAGT TTTTACAAGT AGGCGACAGA TTCTCCATTA AAGAAAAGGA | 14580 |
| AAAAATAAAT GACAAATAAA AATGCATATG CCTCACGTCT CACTACTGAC TAAAGGCTTA | 14640 |
| AGCATTTAGT CAGTAGACGC TTTGTCCTAT AGGATCAAAG CTAGAGCCCT GACTAGTATT | 14700 |
| TTTAGATAAA AAGATGGTTT ATCTAAAAAT ACGTCGCAGT CTTTCTCAAA AAAAGAAAAG | 14760 |
| GAAAAATAAA ATGGCAAATA AAAATGCGTA CGCTCAATCT GGTGTGGATG TTGAAGCGGG | 14820 |
| TTATGAAGTT GTTGAACGGA TAAAAAGCA CGTGGCCCGT ACGGAGCGTG CAGGTGTCAT | 14880 |
| GGGAGCTCTT GGTGGCTTTG GTGGTATGTT TGACCTTTCC AAGACTGGGG TTAAAGAACC | 14940 |
| CGTCTTGATT TCAGGGACTG ACGGTGTCGG AACCAAGCTC ATGTTGGCTA TCAAGTACGA | 15000 |
| CAAGCACGAT ACCATCGGGC AGGACTGTGT GGCCATGTGT GTCAACGACA TCATTGTCTG | 15060 |
| AGGTGCGGAA CCCCTCTATT TTCTCGACTA CGTAGCGACA GGAAGAATG AACCAGCTAA | 15120 |
| GCTAGAACAA GTGGTTGCTG GTGTGGCAGA AGGTTGTGTG CAGGCTGGTG CTGCCCTCAT | 15180 |
| CGGTGGGGAA ACGGCTGAAA TGCCGGGCAT GTACGGCGAA GACGACTATG ACTTGGCTGG | 15240 |
| TTTTGCGGTC GGTGTGGCTG AAAATCTCA AATCATTGAC GGTTCAAAGG TGGTAGAGGG | 15300 |
| AGATGTTCTT CTCGGACTTG CTTCAAGTGG GATTCCTCA AATGGTTACT CTTTGGTTCTG | 15360 |
| TCGTGTCTTT GCGGATTACA CAGGTGAGGA AGTCCTACCA GAATTGGAAG GCAAGAACT | 15420 |
| TAAGGAAGTT CTACTTGAGC CGACTCGTAT CTATGTCAAG GCTGTCTTGC CGCTCATCAA | 15480 |
| AGAAGAGTTG GTCACCGCA TTGCCCACAT CACAGGTGGT GGCTTTATCG AAAATGTCCC | 15540 |
| TCGTATGTTT GCAGATGACC TAGCTGCTGA AATTGATGAA AGTAAAGTTC CAGTGCTTCC | 15600 |
| AATTTTCAAA ACCCTTGAAA AATACGGTCA GATTAAACAC GAAGAAATGT TTGAAATCTT | 15660 |
| CAATATGGGT GTGGGACTTA TGTTGGCGGT CAGCCCTGAA AATGTAGAGC GTGTAAAGA | 15720 |
| ATTGTTGGAT GAAGCAGTCT ATGAAATTGG TCGCATCGTC AAGAAAGAAA ACGAAAGTGT | 15780 |
| CATTATCAAA TGAAAAAAT AGCGGTTTTT GCCTCTGGTA ATGGCTCAAA TTTTCAGGTG | 15840 |
| ATTGCCGAAG AATTTCCAGT GGAGTTTGTG TTTTCAGACC ATCGTGATGC CTATGTGCTT | 15900 |
| GAGCGTGCAA AGCAGCTCGG CGTTCTGTCC TATGCTTTTG AACTCAAGGA GTTTGAGAGC | 15960 |

670

| | | | | | | |
|------------|------------|------------|-------------|------------|------------|-------|
| AAGGCAGACT | ACGAAGCAGC | CCTTGTCGAA | CTCTTGGGAAG | AACACCAGAT | TGACTTGGTT | 16020 |
| TGCCTAGCAG | GCTACATGAA | AATCGTTGGA | CCAACCTTAT | TGTCGGCTTA | TGAAGGTCGG | 16080 |
| ATTGTCAACA | TTTATCCAGC | CTACTTGCCA | GAATTTCCAG | GAGCTCATGG | GATTGAGGAT | 16140 |
| GCTTGAATG | CTGGCGTGGG | TCAGTCTGGT | GTGACCATTG | ACTGGGTGGA | TCGGGTGTG | 16200 |
| GATACAGGCC | AGGTCATCAA | ACAGGTTCTG | GTGCCACGAC | TAGCTGATGA | TACCATTGAC | 16260 |
| AGATTTGAAG | CTCGCATCCA | TGAAGCAGAG | TACAGGCTGT | ATCCGGAAGT | AGTGAAGGCT | 16320 |
| CTATTTACAG | ATTGACTTTT | TGATGATTCA | TATGATATCT | TTGATTTTAA | ATTGGAGTCA | 16380 |
| GTGTTTGTG | AAGACGGCTT | CAAACGGAGG | TATTTGTAAT | GTTAGAATCT | AAAAAACA | 16440 |
| CTCGATATGT | ATTTTATGTC | TATCTGATGT | TATTAACCTG | GGGAATCTTA | TTTAAGTTTG | 16500 |
| AAACAAATCC | TGAATTTATA | GCATTTTCT | TAGCTCCAAG | GTATATCAAT | TGGATTCCAT | 16560 |
| TTTCAGAACC | ACTAATAGTC | GATGGAAAAA | TTGTTTTTGC | TGAAATGTTA | TTTAATCTGA | 16620 |
| TTTTCTTTAT | TCCATTAGGT | GTTTGTTC | CTTTGATAAA | AACTAATTTA | TCTAGTTTAA | 16680 |
| GAATAGTCGG | GACAGGTTTC | TTGATTAGTT | TATTGTTTGA | GTGCTTACAG | TATATTTTAG | 16740 |
| CAATAGGTAT | AACAGATATA | ACGGATTTGA | CTTTAAATAC | GCTAGGTGTC | TGTGTAGGCT | 16800 |
| TACTGATTTA | TCAAATTTTT | ATAAGAGTGT | TCAAATCACA | GACTAGAAAA | TGGATCAATA | 16860 |
| TCTTAGGTAT | GCTTAGCCTT | GGTTTTGCTT | ATCTTGTTTT | ACTGTTACTG | CATTTACTTA | 16920 |
| GTGTTTAACT | AATGATTAAA | AAGGAGAATA | TAATGACTAA | ACGCGTCTTA | ATCAGCGTCT | 16980 |
| CAGACAAAGC | GGGCATTGTT | GAATTTGCC | AAGAACTCAA | AAAACCTGGT | TGGGAGATTA | 17040 |
| TCTCAACAGG | TGGAACAAAG | GTTGCCCTTG | ATAATGCTGG | GGTGGATACC | ATTGCTATCG | 17100 |
| ATGATGTGAC | TGGTTTCCCA | GAAATGATGG | ACGGTCGTGT | GAAGACCCTC | CACCCAAATA | 17160 |
| TCCACGGAGG | GCTTCTCGCT | CGTCGTGACT | TGGATAGCCA | CTTGAAGCG | GCTAAGGACA | 17220 |
| ACAAGATTGA | GCTCATTGAC | CTTGTGGTGG | TCAACCTTTA | CCCATTAAAG | GAACTATCC | 17280 |
| TTAAACCAGA | TGTGACTTAT | GCTGATGCAG | TTGAAAATAT | CGATATTGGT | GGGCCATCTA | 17340 |
| TGCTTCGTTT | AGCAGCGAAA | AATCATGCCA | GTGTTACAGT | TGTGGTAGAT | CCTGCTGACT | 17400 |
| ACGCTGTGGT | TTTGATGAA | TTGGCAGCAA | ACGGCGAAAC | CTCTTATGAA | ACTCGCCAAC | 17460 |
| GTTTAGCAGC | CAAAGTATTT | CGTCACACAG | CGGCTTATGA | CGCCTTGATT | GCAGAATACT | 17520 |
| TCACAGCTCA | AGTGGGTGAA | AGCAAGCCTG | AAAAACTCAC | TTTGACTTAT | GACCTCAAGC | 17580 |
| AACCAATGCG | TTACGGTGAG | AATCCTCAAC | AAGACGCGGA | CTTTTACCAG | AAAGCTTTGC | 17640 |
| CTACAGACTA | CTCCATTGCT | TCAGCCAAAC | AGCTCAACGG | GAAAGAATTG | TCATTTAATA | 17700 |
| ATATCCGTGA | TGCAGATGCT | GCTATCCGTA | TCATCCGTGA | CTTCAAAGAT | AGTCCAACCG | 17760 |

671

| | |
|--|-------|
| TTGTGGCTCT CAAACACATG AATCCATGTG GAATTGGTCA AGCTGATGAC ATCGAGACTG | 17820 |
| CTTGGGACTA CGCTTATGAG TCTGACCCAG TGTCTATCTT TGGTGGGATT GTCGTCCTCA | 17880 |
| ACCGTGAGGT GGATGCTGCG ACAGCTGAGA AGATGCACGG CGTTTTCCCTC GAAATCATCA | 17940 |
| TTGCACCAAG CTATACGGAT GAAGCGCTAG CCATTTTGAT CAATAAAAAG AAAAAGTTGC | 18000 |
| GTATCCTTGC CTTGCCATTT AATGCTCAAG AGGCTAGCGA AGTGAAGCA GAATACACAG | 18060 |
| GTGTAGTCGG TGGACTTCTC GTGCAAAATC AAGACGTGGT CAAGGAAAGC CCAGCTGACT | 18120 |
| GGCAAGTGGT GACTAAACGT CAGCCAACTG AGACAGAAGC GACTGCTCTT GAGTTCGCTT | 18180 |
| GGAAGGCTAT CAAGTACGTC AAATCAAATG GTATTATCGT GACCAACGAC CACATGACAC | 18240 |
| TTGGTGTGG TCCAGGTCAA ACCAACCGTG TGGCTTCTGT TCGCCTTGCC ATTGACCAAG | 18300 |
| CCAAAGATCG TCTGGACGGG GCGGTCCTTG CTTGAGATGC CTTCTTCCCA TTTGCGGATA | 18360 |
| ACGTGGAAGA AATCGCCAAA GCAGGAATTA AGGCCATCAT CCAGCCCGGT GGCTCTGTCC | 18420 |
| GTGACCAAGA ATCCATCGAA GCAGCGGATA AATACGGCTT GACTATGGTC TTTACAGGTG | 18480 |
| TGAGACATTT TAGACATTAA GAAGATAAAA GGGAAGAAAA CAGTTTCTTT CCTTTTGTGG | 18540 |
| CTTAAATAC TAACTGAAAC AAGATTAAAA CGAATTTTTT TGATATAATG TTGGTAAATA | 18600 |
| ATTCGCAAAA GAGGTGAGG AATGAAACTG CTTGTTGTGCG GTTCTGGTGG TCGTGAGCAT | 18660 |
| GCGATTGCTA AAAAGTTACT TGAATCAAAA GACGTGGAAG AAGTCTTTGT AGCTCCTGGG | 18720 |
| AATGATGGGA TGACTCTGGA TGGTTTGGA TTGGTAAATA TCTCTATTTT CGAACATTAT | 18780 |
| AAATTGATTG ACTTCGAAA GACCAATGAT GTTGCTTGGA CCTTTATCGG TCCAGATGAT | 18840 |
| GCCCTTGCTG CTGGTATCGT GGATGATTTT AACCAAGCTG GACTTAAGGC CTTTGGTCCG | 18900 |
| ACTAGGGCTG CAGCGGAGCT GGAGTGGTCC AAGGATTTTC CCAAGGAAAT CATGGTCAAA | 18960 |
| TACGGCGTTC CGACAGCAAC ATATGGCACA TTTTCAGATT TCGAGGAAGC CAAAGCCTAT | 19020 |
| ATCGAAAAGC ATGGTGCGCC TATCGTAGTC AAGGCGGATG GCTTGGCACT TGGGAAGGGT | 19080 |
| GTCGTCGTTG CTGAGACGGT TGAGCAAGCG GTCGAAGCCG CTCATGAGAT GCTTTTGGAC | 19140 |
| AATAAATTTG GTGACTCAGG TGCGCGCGTG GTTATTGAGG AATTCCTTGA AGGAGAGGAA | 19200 |
| TTTCACTCT TTGCCTTTGT CAATGGTGAT AAGTTCTACA TCATGCCAAC GGCTCAGGAC | 19260 |
| CACAAACGTG CCTATGATGG CGACAAAGGG CCTAACACGG GTGGTATGGG TGCCTATGCG | 19320 |
| CCAGTCCCAC ACTTACCACA GAGTGTAGTT GATACAGCGG TTGACACCAT TGTCAAGCCA | 19380 |
| GTCTAGAAG GGGTGATTAA AGAAGGTCGC CCTTATCTGG GAGTTCTTTA CGCAGGGCTT | 19440 |
| ATCCTGACAG CTGATGGACC GAAAGTCATT GAGTCAACG CTCGGTTCGG AGATCCAGAA | 19500 |

672

| | |
|--|-------|
| ACTCAGATTA TCTTGCCTCG CTTGACCTCT GACTTTGCTC AAAATATCAC AGATATCCTG | 19560 |
| GATAGCAAGG AGCCAAATAT CATGTGGACG GACAAGGGTG TGA CTCTGGG TGTGGTTGTC | 19620 |
| GCATCCAAGG GCTACCCGCT AGACTATGAA AGGGGCGTTG AGTTGCCAGC CAAGACAGAA | 19680 |
| GGCGATGTCA TCACCTACTA TGCAGGGGCT AAGTTTGCGG AAAATAGCAG AGCACTGCTC | 19740 |
| TCAAACGGCG GACGAGTTTA TATGCTCGTT ACCACAGCAG ATACCGTCAA AGAAGCCCAA | 19800 |
| GCCAGCATAT ACCAAGAACT ATACCAACAA AAAATAGAAG GACTCTTCTA CCGAACAGAT | 19860 |
| ATCGGAAGCA AGGCAATTAA GTAAAGATAT AAGAATAACG CGCCGTAGTC GCCAAACACG | 19920 |
| ATAATGGTCG TCGTGGTGAA AAGACCAGAA CAGTGAATGT TCTGGTCAGG GGGAACTTG | 19980 |
| GAGACCTTAG GCTCAAAGTT TAGGAATGAA ACCGAAGGTT TGCTTCCGCC TCCATCACCT | 20040 |
| AAGACCATTA TCAAAAAGAA AAATAAAAAT TCACAAAATA CGTTAATGAT CGTATGGTTT | 20100 |
| GCGAGCGTTA GCGAGCTAAT ATAGAACAAT CACCGCCGTT GTGAAAGAAC GATTGGATGA | 20160 |
| TAATCCAATC GTTCAGGGAA ATTGGAAGAC CTTGGGTTTC CAATTTAGGC ATGAGACACC | 20220 |
| TTTGGTGGCT GCTGCCGTCC CTCACAAGCT AAGGTGATTG TTGAAAAAGA GGAAAAAGGA | 20280 |
| GAAGAAATGA AACCAGTAAT TTCCATCATC ATGGGCTCAA AATCCGACTG GGCAACCATG | 20340 |
| CAAAAAACAG CAGAAGTCCT AGACCGCTTC GGTGTAGCCT ACGAAAAGAA AGTTGTTTCC | 20400 |
| GCACACCGTA CACCAGACCT CATGTTCAAA CATGCAGAAG AAGCCCGTAG TCGTGGCATC | 20460 |
| AAGATCATCA TCGCAGGTGC TGGTGGCGCA GCGCATTTGC CAGGCATGGT AGCTGCCAAA | 20520 |
| ACAACCTTC CAGTCATTGG TGTGCCAGTC AAGTCTCGTG CTCTTAGTGG AGTGGATTCA | 20580 |
| CTCTATTCTA TCGTTCAGAT GCCGGTGGG GTGCCTGTTG CGACCATGGC TATCGGTGAA | 20640 |
| GCTGGAGCGA CTAACGCAGC TCTCTTTGCC CTCCGTCTCC TCTCTGTAGA AGATAAGTCC | 20700 |
| ATTGCGGATG CACTTGCCAA CTTTGCTGAA GAACAAGGAA AAATCGCAGA GGAGTCGTCA | 20760 |
| AATGAGCTCA TCTAAAACAA TCGGAATTAT CGGTGGCGGT CAACTGGGTC AGATGATGGC | 20820 |
| CATTTCTGCT ATCTACATGG GCCACAAGGT TATCGCGCTG GATCCTGCGG CGGATTGCCC | 20880 |
| GGTCTCTCGT GTGGCGGAAA TCATTGTGGC ACCTTATAAC GATGTAGACG CCCTCCGTCA | 20940 |
| GTTGGCAGAC CGTTGCGATG TCCTCACTTA TGAGTTTGAA AATGTCGACG CTGACGGTTT | 21000 |
| GGATGCCGTT ATCAAGGATG GACAACTCCC TCAAGGAACA GATCTGCTCC GCATTTGCA | 21060 |
| AAATCGTATT TTTGAAAAGG ACTTTTGTG AAACAAGGCT CAAGTCACTG TGGCACCCTA | 21120 |
| CAAGGTCGTG ACTTCTAGCC TAGACTTGGC AGATATCGAC TTGTCGAAAA ACTATGTCCT | 21180 |
| CAAGACTGCG ACTGGTGGCT ACGATGGTCA TGGCAAAAAG GTTATTCGTT CAGAAGCAGA | 21240 |
| CTTGGAAGCA GCCTATGCGC TAGCAGACTC AGCAGACTGC GTCTTGAAG AATTGTCAA | 21300 |

673

| | |
|---|-------|
| CTTTGACCTT GAGATTTCTG TCATCGTGTC AGGAAATGGC AAGGAGGTGA CGTTTTTCCC | 21360 |
| AGTTCAGGAA AATATCCACC GCAACAATAT CCTGTCTAAG ACCATCGTAC CAGCCCGCAT | 21420 |
| TTCTGAAAGT CTAGTAGACA AGGCTAAAGC TATGGCAGTG CGAATCGCAG AACAACTCAA | 21480 |
| CTTGTCTGGA ACTCTCTGTG TGGAAATGTT TGGCAGAGCT GATGACATCA TTGTCAATGA | 21540 |
| AATCGCCCCA CGACCACATA ACTCTGGGCA CTATTCTATT GAAGCCTGTG ATTTCTCTCA | 21600 |
| GTTTGACACC CATATTCTGG GTGTTCTCGG AGCACCATTA CCAGTCATCA AACTCCATGC | 21660 |
| GCCAGCCGTT ATGCTTAATG TCCTCGGTCA GCATGTCGAG GCTGCTGAAA AATATGTCAC | 21720 |
| AGAAAATCCA AGCGCCACCC TCCACATGTA TGGTAAAATA GAAGCAAAGC ATAATCGTAA | 21780 |
| GATGGGACAT GTGACTTTGT TTAGTGATGT GCCGGATAGT GTGGAAGAGT TTGGGGAAGG | 21840 |
| GATTGATTTT TAGGACAAGT CTATGATACA AATTATCGTT AATACATTTA TTGAAAAGTA | 21900 |
| TAAGACTGGA GCAGTTGTTG AAGTGTTGTA TGCCAGTGCT GACCAAGATA AGGTACAAGC | 21960 |
| TAAATATGAA GAACTAGCTG CACAATACCC CGAAAATTAT TTAGCTATCT ATAATGTACC | 22020 |
| GCTGGATACG GATTTGAATA CACTAGATCA TTACCCGTCT GTGTTTATTG GAAAAGAGGA | 22080 |
| GTTTGAGTAG AAATCTTGGT TTACCTAGAT AGCTTATTCC CAACAGCTTA AGAAGAAAGG | 22140 |
| AAAAATTAAC ACATGATCAA CCGTTACTCT CGCCCTGAGA TGGCGAATAT TTGGAGTGAA | 22200 |
| GAAAATAAAT ACCGTGCTTG GCTTGAGGTG GAAATCCTCT CTGACGAGGC ATGGGCTGAG | 22260 |
| TTGGGGGAAA TCCCTAAGGA AGATGTGGCT TTGATTCGCA AGAAGGCGGA CTTTGACATC | 22320 |
| GACCGTATTT TGGAAATGA GCAGGAGACG CGCCACGATG TGGTGGCTTT CACGCGTGCG | 22380 |
| GTTTCTGAGA CTCTTGGTGA AGAGCGCAAG TGGGTTCAC TGGGTTAAC TTCTACTGAC | 22440 |
| GTGGTGGATA CTGCTTATGG TTACCTCTAC AAGCAGGCCA ACGACATCAT CCGTCGTGAC | 22500 |
| CTTGAAAAC TCACTAATAT CATCGCTGAC AAGGCCAAGG AGCACAAGTT CACCATCATG | 22560 |
| ATGGGGCGTA CTCATGGTGT GCACGCTGAG CCGACAACCT TTGGTCTTAA ATTAGCAACT | 22620 |
| TGGTACAGCG AAATGAAACG CAATATCGAG CGCTTCGAGC ATGCGGCTGC TGGTGTAGAA | 22680 |
| GCTGGTAAGA TTTCTGGTGC GGTGGGAAC TTTGCCAATA TCCCACCATT TG TAGAGGAG | 22740 |
| TATGTCTGCG ATAAACTTGG CATCCGTGCC CAAGAAATCT CTACACAAGT CCTTCCTCGT | 22800 |
| GACCTTCACG CTGAGTACTT TGCGGTTCTT GCCAGCATTG CGACTTCAAT CGAACGTATG | 22860 |
| GCGACTGAGA TTCGTGGTCT AAAAAATCT GAGCAACGCG AAGTAGAAGA GTTCTTTGCT | 22920 |
| AAAGGGCAAA AAGGGTCTTC AGCAATGCCT CACAAACGCA ACCCAATCGG TTCTGAAAA | 22980 |
| ATGACTGGTC TGGCGCGTGT CATTCGTGGT CACATGATTA CGGCTTATGA AAACGTCGCT | 23040 |

674

| | |
|---|-------|
| CTCTGGCATG AACGCGATAT TTCTCACTCA TCAGCTGAGC GTATCATCAC ACCAGATACG | 23100 |
| ACCATTTTGA TTGACTACAT GCTCAACCGT TTTGGAAATA TCGTCAAGAA CTTGACAGTC | 23160 |
| TTCCCAGAAA ATATGATCCG AAACATGAAC TCGACTTTTG GTCTTATCTT TAGCCAACGG | 23220 |
| GCTATGTTGA CATTGATTGA AAAAGGCATG ACCCGTGAGC AAGCCTATGA CTTGGTGCAA | 23280 |
| CAAAAACAGC CTACTCTTGG GACAACCAAG TAGACTTTAA ACCACTTCTT GAGGCAGATT | 23340 |
| CAGAAGTAAC ATCACGTCTC ACACAAGAAG AAATCGATGA AATCTTCAAC CCAGTTTATT | 23400 |
| ACACCAAACG AGTGGATGAT ATCTTTGAAC GTCTTGGACT AGGTGATTAA TTAATAAATA | 23460 |
| AACAGCGAGC TTCAATCTCG CTGTTTATTT TTTATCGAAA AGACTTAGTC TTCTTTTCTT | 23520 |
| TTAGTGAGTC CATAGGCTGC TAGTGTGGAC ATGAGTCCTG CGACTACTAG TCCTGCAGAA | 23580 |
| TCGTGAGTTC CTGTTTCAGG AAGTTTTTTC TCTGTTACCA CAGGAGCTGG ATCTTGAGGA | 23640 |
| AGAACTTTGC TTTCCTCAGC AGGAGCAGTT GATGGAGCTG GTTGGCTTGG GATTTCTAGT | 23700 |
| TTTGGTTTTT CTTCAGCAAT AGCGGCTTGT CCGTTTTTCAT CGCCTACATG TGTACCATA | 23760 |
| GTTCGCACTT CGACTATTTG AGTAACGGCT TCCTGTGCTA CGACACTATT TACAAGTGTT | 23820 |
| TTCACTTCCT TACCATCGGC AGAAGTGCTC ACAGAGTAGA AGTTGCTACG ATGTCCATTG | 23880 |
| ACGCCCTTAG TAATGACTTG TGTTTTTCCT TTGAGTAAGA GTGGATTTTC ACAAGTCACT | 23940 |
| GTGGTAAATG GAATTTCTTC TTCTTGATA TCCAGTCTAG GTTTTACCTC AGTAGTTGGT | 24000 |
| GCAAGACCAC TTTCATCACC CTTGTGAGTT ACAGGAGCGC CAACTTCAAC CACTTGGTTT | 24060 |
| ATAACTTCTT TGGTTACCTG GCTATCAAGG ACTGTTTCTG TTGTTTTTCC ATTTTCAGTG | 24120 |
| AGTACAGAGA TGTAATGAGT TCGTTCACCT TTGACTCCTG CTGTGATAAT ATTTTCCTGA | 24180 |
| CCGGCTGGGA GGTTAGGATT TTCTTTCTTG ATAACCTCAA ATGGAATTTT TTCAGTTCTT | 24240 |
| GTGATGAGTT CTGGTCTGGT TTCAACATTG GCAGCCACTT CATTTTCATC TAGGCTTCCT | 24300 |
| GAATGAGTTA CAGCTGGTTT GAGGCCTTGA AGAGCGGCTT TTAGGTTGGC TACAAGCGTG | 24360 |
| TCAAGCTCAG CTTGTTTATT ACGGTTGAGG TTGTAATTGA GAGCTGTTTT AGCTGCGTCA | 24420 |
| AGGGCCTCAA GACTTTCTTT ACTATATCCT TCTAAGTTTG TAGGAATTTT AGCTAATTCT | 24480 |
| TCGCGGAGAG CATTATAATT AGCACGAAAG TAGTCTTTGT TGTGGTCTGC AAAGGCAGTC | 24540 |
| ATGAGTTCAA AGATTTCTTC TTCCTTGAT TCAGCGCTTG GTCTATCTGC CCAGATTGAA | 24600 |
| AGCATACTTC CGACTGTTGG AAGATCTACT TCAGGATATT TGGTAGAAGC TAGTTGATTG | 24660 |
| AATGGTGTTT TTCCAGTATT CTCAATAGCT TTCTTGAGGA AACCACCACC ATCTTCTGGT | 24720 |
| TTTTGACCAA GAATGTAGTA CCAGTCACCG TTGGTATTCA AGAATTTATA GCCTTTGCTT | 24780 |
| GCTAGGTATT GAGGTGATGC GAGGTTATAT CCCACCAGC CTTTAGACCA GTAAGAAATC | 24840 |

675

| | |
|---|-------|
| AAGACATCTT TGTCAAACG AACATCGTCC TTGTCTTCAT AGTAGAAGCC ATCGTTGAAG | 24900 |
| GCCATTGGTT GAAGCCCTCT TTCTTTGGCC ATAGCTGCCA GGGTGTGGC ATATTCGGCA | 24960 |
| AATTTGCCAT AGAGTTGATA CCACTTGAGG TAGTACCAGC CTTGGGCACT AGTCGCATCG | 25020 |
| TTGGCGTATT CGTCAGTACC AAAGTTGAAA ATCTTTGTTT TACCTGCAAA GAAGTCCATG | 25080 |
| TATTTACCGA TGAGGGCTTT TACAAAGTTC ATCGCTTCTT CGTTTTTCAA GTCCATAGTT | 25140 |
| ·GTTTTTGAAA CTTTATCAAA GTGGGCTTGA GGATTTTAA TACCTAATTT TTCCATGGCA | 25200 |
| ACCAGCATAG CATCCATGTG ACCTGGACTG TTAATAGCTG GGATGAGACC GATGTCTTA | 25260 |
| GATTTAGCGT ATTCAATTAG CTCTGTTACT TCTGCCTGTG TTAGTGCACT ACCGTTTGGA | 25320 |
| TCGTCGTAGT AAGCTTTAGT TCCTTCGATA ATAGCTTTTT TAACGTCATC ACTAGCATAG | 25380 |
| GTTTTTCCGT TGGCAGTAAT GGTCATATCA TCGAGTAGAA AGCGAAGTCC GTCATTTCTT | 25440 |
| AGAAGGAGAT GGACATCAGA ATATCCGAGC TCACTGGCCT TGTCTACGAT GCGTTTGAGC | 25500 |
| TGGTTCAGAG TAAAGTATTT GCGTCCAGCA TCGATTGAGA TTACCTTGTT TTTGGCAAGT | 25560 |
| TTTTCAACCT CACGTTAGC TTCTTCTTCT TTTTGAGCTT CAGGCGTGAG GGTCAAGTTG | 25620 |
| TTGACAGTTT CTTGAAGTTT AGCAATGGCT TGATCAATCG TATCTTGTTG GGCACGGCTA | 25680 |
| AGGTTGCTAT CGAGAGAGCG AATAGCTTTT TCAGCTTCTT TTACGGCCGT GACGCTTCTT | 25740 |
| GCAGTATAAC GGTTCAGGTC TTTTGGTACC TCGTTAAGTG CTTGCTCTGC AGATTCTATA | 25800 |
| TCAGCTGCGA AGTATTCAGC GTTGGCATTT GCAAATGAC GCATGAGTTT GAAGAGGCGT | 25860 |
| GATGGTGAAT AACGTGCAGA TGGAGTGTCA GCCCAAGCAG CTACCATACC ACCGATGATT | 25920 |
| GGGATATCAG CTCCTTCTGT TTTTGGTACA GAAGTGATTG GTGTGTTTTT AATACCATTG | 25980 |
| AGCCCCTGAT CGAGATTGTA CCAGCCTTGG CCATCAGCGT TTCGTCCAAG AACGTAGTAC | 26040 |
| CAAGCATCAT TGGTATTAAG GATTGGTGA CCTTTTTCAG CTAGTAGTTT AGAAGAAGCG | 26100 |
| ACATCGTAGC CTCCCCAACC ACCAGTCCAC ATAGAAACGA TGATGTCTTT GTCAAAACTA | 26160 |
| CCAAAGCTTG TGTCGCTATT GTAGTAGATA CCGTCGTAA AAGCCATTGG TTTGAGACCG | 26220 |
| TGCGATTTTA CAATACGAGC GAGGTCATTG GCGTAGGCAA TAAATTTTTC ATAGCCTTTT | 26280 |
| ACAGGGTAGC CTTGTTTGG ATAGTATTTA TCAGCTTGAA GCACACTCCA ACCTTTAGCA | 26340 |
| TCTGTGCGAT CATTGGCATA TTCATCAAGT CCGATGTTGA AGATTTCACT CTTTTTCGCG | 26400 |
| AAATAAGCAG CATACTTGTC GATAAGGGCT TTTGTAAAAG CGACAGCTTG TTCGTTGTCA | 26460 |
| AGATCGACAG TACGGGCTGA TTCTTCCCA AAATAGCTAA AGTTAGGGTT TTGGATTCCC | 26520 |
| AATTCTTTCA TGGCATTGAG AATCGCATCC ATGTGTCCAG GACTATTTAC TGTCGGAATG | 26580 |

676

| | |
|--|-------|
| AGACCGATAC CTTTATCTTT GGCATAGTTA ATCAGATCTG TCATTGACT TTCTGTTAAG | 26640 |
| TGATTGCCGT TTGGATCGTT GTAATAATCA TTTGTACCTT TTTCAATGGC GCGTTTGACA | 26700 |
| TCGTCACTGG CATAGGTCTT GCCGTTAGCT GTGATGCTCA TATCGTCCAA CATGAAACGG | 26760 |
| AGTCCATCAT TTCCGACTAA TAGGTGTAAG TCAGTGTAGC CATAATGTTT CGCTTTATCG | 26820 |
| ATGATTTCCCT TGAGCTGTTT TGGTGAGAAA TATTTACGTC CAGCATCAAT AGAAACAATT | 26880 |
| TTCTTTTTCG CTAGTTTTTC ATTTACAGTT GCAGCACGTT CCTTTCCTGC CTCTGTTGCC | 26940 |
| GGTTTGTCAG CCTCTGCTTT CGCTTCATCT TTTTTAGCTG GTTTATCCTT GTCAGTCTTG | 27000 |
| TCTGTATTG ACTCTTTAGA ATCAACCTCT TTCGCTTCTT CCTTTTTAGG GCTAGCTTCT | 27060 |
| TCTGCCTTTT TATTAGCAGT TTCTTTTCA GCAGAAGTTG GAGTTACCAC TTCTGCTTTA | 27120 |
| TCACTAGGAG TTGAACAAAC TTCCTCTGT GGTTTTTCTT CTGTTTTTGG AAGACTAGCT | 27180 |
| ACCTTATCAG TAGCTGGAGT TTCTGTTTCT ACAGTTTTTG GAGCTTCTGG TTGAAGCACT | 27240 |
| GCTTTAGGTG TTTCCTCAGT CCGATTTTCG GATGATTGAG GGAATCAGA AACCGTATGG | 27300 |
| ATGGTCGGTT GGTTTTCTGT AGTAGTAGA GTAACCTCAT CGGCTGCAAC AGTCTGTGCT | 27360 |
| TGGAAGGCAA ATCCAATTAG AACAGAAGCT GCTCCTACAG CGTATTTACG AATAGAAAAA | 27420 |
| CGCTGTGTT TTTTATGTTT CATTGCAAAA CCTCCTGATT GCATTGTTAT ATTGATAGCG | 27480 |
| ATTATATAAA TCAACGCCTT TATTTTATTT CTTATATTAA TTTCTTATAT TAACGAGAGT | 27540 |
| CAAGAGGAGA TGACAAAAA CTATAATAAG TATAAAAAA TATAAAATTT AAACCTAAGA | 27600 |
| TTTCAGATTG GTCGAAAAA ATACGTATAT ATATCTAGTA TAATTTTGG TTCTATTCT | 27660 |
| ATAAAATATT CCACAAATTA TAGAATTTT CAAAATAGG TAAGCGCTAC CTTTTTGGTG | 27720 |
| TAGTATAATA AGCATAGAAA AAGCCCAAGC GATTAGCTCA GGTTTTCTTC TTAGTGATCA | 27780 |
| CGGTACATG AGATAAATTT AATCTTGTAG TAATCAGATC GTTTGTAAGT TTCACTGTAT | 27840 |
| TCTAAACTT GCCCAGTTGA TTCGAGTTG GTGATTTTAG TTTGTAGGAC AGTAGGGAAT | 27900 |
| TGTTTCATCGA CTCCGAGGAC TGAAGCTGCA TGTCTGGAG TTGGAAGAC TATTTCTGTG | 27960 |
| ATTTCTTCAA AGTGTTTATC ATTCATGTGA ATGTGGTAGT CTAACCTGAA ACGATTATAG | 28020 |
| ATAGAACTAT AGTATTCAAG GTTTGGATAA TTTGCGTTGA TATATGTTT TGGGATGTAG | 28080 |
| GATGTATGGT AGATATAAAC GACACCGTTT GATTCGCGGA TACGTTCAAT CTTGTAGTAG | 28140 |
| AATTGATCGC CGCGTAGACC CAATTTTCC AAGTAAACAA GCTTGTTCCT GCGTTCAATT | 28200 |
| GAAAGAACAG TTACCTTATC ATCTTTAGCA TTGAAGAGTT CAATATCTGA AAACCTTACA | 28260 |
| AGCTTGTGTT TGCGTGACG TGAAACGAAG GTTCCTTTTC CTTGTTGGCG GACAATATAG | 28320 |
| CCATCTTTGG CAAGGTCGTT TAAGGCGCGA ACAACTGTGA TAGAGCTGAC ATCGTACATT | 28380 |

677

GAAATGAGTT CTGCTTCAGT GTAAAATTTA TCTCCACTGC TAAACTGCCC AGAGATGATT 28440

TTATTTTTTA ATTCGTCTTT TATGTATTGA TGG 28473

(2) INFORMATION FOR SEQ ID NO: 84:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6749 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 84:

| | |
|--|------|
| CCTGATGGGT GGTATGCGAG GATACAGTTC TGAAAATCGC CGTTACTTAA TTAATGGACG | 60 |
| CGAAGTCACA CCTGAGGAAT TTGCTCACTA TCGTGCGACT GGTC AATTAC CAGGAAATGC | 120 |
| AGAAACTGAT GTGCAAATGC CACAACAGGC ATCAGGTATG AAACAAGGCG GTGTCCTTGC | 180 |
| AAAAGTAGGT CGAAACTTAA CAGCAGAAGC GCGTGAGGGC AAGTTGGATC CTGTTATCGG | 240 |
| ACGAAACAAG GAAATCAAG AAACATCTGA AATCCTCTCA CGCCGCACCA AGAACAATCC | 300 |
| TGTTTTGGTC GGAGATGCAG GTGTTGGTAA GACAGCAGTT GTCGAAGGTC TAGCGCAAGC | 360 |
| CATTGTGAAC GGAGATGTTT CTGCTGCTAT CAAGAACAAG GAAATTATTT CTATTGATAT | 420 |
| CTCAGGTCTT GAGGCTGGTA CTCAATACCG TGGTAGCTTT GAAGAAAATG TCCAAAACCTT | 480 |
| AGTCAATGAA GTGAAAGAAG CAGGGAATAT TATCCTCTTC TTTGATGAAA TTCACCAAT | 540 |
| TCTTGGTGCT GGTAGCACTG GTGGAGACAG TGGTTCTAAA GGAAGTTCGG ATATTCTCAA | 600 |
| GCCAGCTCTC TCTCGTGGAG AATTGACAGT GATTGGGGCA ACAACTCAAG ACGAATACCG | 660 |
| TAACACCATC TTGAAGAATG CTGCTCTTGC TCGTCGTTTC AACGAAGTGA AGGTCAATGC | 720 |
| TCCTTCGGCA GAGAATACTT TTAATAATCT TCAAGGAATT CGTGACCTCT ATCAACAACA | 780 |
| CCACAATGTC ATCTTGCCAG ACGAAGTCTT GAAAGCAGCG GTGGATTATT CTGTTCAATA | 840 |
| CATTCCCTCAA CGTAGCTTGC CAGATAAGGC TATTGACCTT GTCGATGTAA CGGCTGCTCA | 900 |
| CTTGCGGGCT CAACATCCAG TAACAGATGT GCATGCTGTT GAACGAGAAA TCGAAACGGA | 960 |
| AAAAGACAAG CAAGAAAAAG CAGTTGAAGC AGAAGATTTT GAAGCAGCTC TAAACTATAA | 1020 |
| AACACGCATT GCAGAATTGG AAAGGAAAAT CGAAAACCAC ACAGAAGATA TGAAAGTGAC | 1080 |
| TGCAAGTGTC AACGATGTGG CTGAATCTGT GGAACGAATG ACAGGTATCC CAGTATCGCA | 1140 |
| AATGGAAGCT TCAGATATCG AACGTTTGAA AGATATGGCT CATCGCTTGC AAGACAAGGT | 1200 |
| GATTGGTCAA GATAAGGCCG TAGAAGTTGT AGCTCGTGCT ATCCGTCGTA ACCGTGCTGG | 1260 |

678

| | |
|--|------|
| TTTTGATGAA GGAAATCGCC CAATCGGCAA CTTCTCTCTT GTAGGGTCTA CTGGGGTTGG | 1320 |
| TAAGACGGAG CTTGCTAAGC AATTGGCACT CGATATGTTT GGAACCCAGG ATGCGATTAT | 1380 |
| CCGTTTAGAT ATGTCTGAAT ACAGTGACCG CACAGCTGTT TCTAAGCTAA TTGGTACAAC | 1440 |
| AGCAGGCTAT GTGGGTATG ATGACAATAG CAATACCTTA ACAGAACGTG TTCGTCGCAA | 1500 |
| TCCATACTCT ATCATTTCTT TGGATGAAAT TGAAAAGGCT GACCCCTAAG TTATTACCCCT | 1560 |
| TCTCCTCCAA GTTCTAGATG ATGGTCGTTT GACAGATGGT CAAGGAAATA CAGTAAACTT | 1620 |
| CAAGAACACT GTCATTATTG CGACCTCAA TGCTGGATTT GGCTATGAAG CCAACTTGAC | 1680 |
| AGAAGATGCG GATAAACCAG AATTGATGGA CCGTTTGAAA CCCTTCTTCC GTCCAGAATT | 1740 |
| CCTCAACCGC TTTAATGCAG TCATCGAGTT CTCACACTTG ACTAAGGAAG ACCTTTCTAA | 1800 |
| GATTGTAGAT TTGATGTTGG CTGAAGTTAA CCAAACCTTG GCTAAGAAAG ACATTGACTT | 1860 |
| GGTAGTCAGT CAAGCGGCTA AAGATTATAT CACAGAAGAA GGTACGACG AAGTCATGGG | 1920 |
| GGTTCGTCCT CTCCGTCGCG TGGTTGAACA AGAAATTCGT GATAAGGTGA CAGACTTCCA | 1980 |
| CTTGATCAT TTAGATGCTA AACATCTGGA AGCAGATATG GAAGATGGCG TTTTGGTTAT | 2040 |
| TCGTGAGAAA GTCTAAGACA GAATTTTGAG GATAAAAAAG AAGGAGCCAG CTGAAAAAA | 2100 |
| CTGGTTCCCT TTTAGGTACG ACAGGCATGT CGTATAGTAG AAGTGATTA TTCTAGTTTC | 2160 |
| AATATACTAT AGTAGCTCAG AAGTCGGTAC TTAAACGTGC TATATCAAAA CCAGTCCTGG | 2220 |
| AAAAACGTGG ACTGGTTTCG TGTTTGGATT ATTACCTTGA ACGACATGCG TTAAGTTA | 2280 |
| GTGAACCGC CGTATGCCGA ATGGTACGTA CGGTGGTGTG AGAGGGGCTA GAGATTATCC | 2340 |
| CCTACTCGAT TTAAATCAC ATGACGTTCA AAGGCATCAT CTGAAATCCC TTGTTCCAAG | 2400 |
| ATGAGTTTGG CCCATTCTTT AGCAGAGAAG AGGCTGTGGT CCTGTAGTT TCCGCAAGAT | 2460 |
| TCGATGGTTG TCCCTGGGAC ATCTTCCCAA GTAGTAGTTT CAGCGATTTC CTGAGCGAA | 2520 |
| TCCTTGATAA CAGCTCGCAT TTTAGCACTG GTGTGACGTC CCCACATAAT CATGTGGAAG | 2580 |
| CCTGTGCGGC AACCAAATGG TGAACAGTCA ATCATGCCGT CAATGCGGGT ACGGATGAGT | 2640 |
| TTGGCTAAGA GGTGCTCGAT AGTGTGAAGG CCGGCAGTAG GGATAGAGTC TTCGTTTGGT | 2700 |
| TGCACCAAGC GAATATCATA ATTGGAGATG ATGTCTCCTT TTGGTCCGT TTCTTCCCCA | 2760 |
| ATCAAGCGAA CATAGGGTGC TTTGACAATG GTGTGGTCAA GTTCAAACT TTCGACAATA | 2820 |
| ACTTCTTTTG ACATGGTAAA TCCTTTCAGT TTTCTTCTCT CATTATATCA TAAAGGTTGC | 2880 |
| TCCTGAGACA GAGAGAAAAC CTCTCCGAGG CTGGAGAGGT TGAAATCTTT ACTTACGATA | 2940 |
| TAAGCGGTCG TATTGGTAGT ATGGGTCAA GGTACGTTG ATACCCAGTT TACGAAGGAC | 3000 |
| ATTCTTGTCT TCATCAGTCA AGATGATGGT TGAGTGGGCT TCGCTTCCTT TGAGGTTGCC | 3060 |

679

| | |
|--|------|
| GAGTTCCTCC ATAGCGCGGG CAGCATCAGG ATTTCTGTGTA GCTGTGATAG CAAGTGCAAT | 3120 |
| CAGGATTTC A TTTGAATGAA GCGGTGGATT GCGGCTACCG AGATGATCGA TTTTAAGACC | 3180 |
| TTGGATTGGC TTAACAACTT CAGGCTCGAT TAGTTTACT TCTTTAGCGA TGTCAGCTGA | 3240 |
| TTTTTTGATG GCGTTGATCA AGGCAGCGGC TGTAGGACCA AAGAGTTCTG AGTTCTTACC | 3300 |
| AGTGATGATT TCCCCATTTG GCAATTCAAA GGCTAGGGCT GGTCCACCAG TTTCTTCTGC | 3360 |
| TTTTTTGGCGC GCAACGACAG CAACCTTACG GTCTGCAGGT GTGATACCGA GGTCTTCAT | 3420 |
| GAGCAACTCA ATTTTCTTGA CGGCAGCTTC GCCAACTTTT TCAGCTTTGA AGTCAAGAAC | 3480 |
| TGTTTGATAG TAACGCGGGA TGATTTCTTG TTTAGAAGCT TCGACAGCGG CCTCGTCATC | 3540 |
| TGTAATAGCG AAACCAACCA TGTTGACACC CATATCTGTC GGTGAAGCGT ATGGTGATTT | 3600 |
| TCCGAGAATA CGTTCCAACA TGCGTTTGAG CACTGGGAAG ATTTTCGATAT CACGGTTGTA | 3660 |
| GTTGACAGTG GTTTCTCCAT AGGTTTGAAG ATGGAAGGGG TCAATCATGT TGACATCATC | 3720 |
| AAGGTCAGCT GTGGCAGCTT CATAAGCCAA GTTAACTGGA TGATGAAGGG GAAGATTCCA | 3780 |
| AACAGGGAAG GTTTCAAAT TAGCGTAGCC AGATTGATG CCATTGATTT GGTCGTGGTA | 3840 |
| CATATTGGAC ATACACGTTG CCAATTTTCC AGAACCAGGT CCAGGAGCGG TTACGACAAT | 3900 |
| CAAGTTGCGA CTGGTTTTGA TGTAGTCGTT TTTGCCCATG CCTTCTGGGG AAATGATGTG | 3960 |
| ATCCATATCC GTCGGATATC CTTTGATTGG ATAATGAAGA TAAGAATCAA TTCCGTTTTT | 4020 |
| CTCAAGTTGA TTGCGGAAG CATCTGCAGC GGGTTGGCCA GCGTATTGTG TAATGACAAC | 4080 |
| GGAACCAACA AAAATCCCTA ATTCAATGAA TTTATCAATC AAACGAAGAA CTTCTTGGTC | 4140 |
| ATAAGAAATG CCTAAGTCGC CACGTGCTTT GGAATGTTCA ATGTTGCTAG CATTAATGGC | 4200 |
| AATCACAACC TCAACCTGCT CTTTCAATTC TTGCAAGAGC TTGATTTTGT TGTCAGGTTT | 4260 |
| ATAACCAGGA AGGACACGAG CAGCGTGGAA ATCTTCTAAC ATTTTACCGC CAAACTCTAA | 4320 |
| GTAGAGCTTG CCGTCAAATT GGTTAATGCG CTCGAAAATA TGGTCGCGTT GTAAATTCAA | 4380 |
| ATATTGTTCA GAACTAAAAG CTTGTTTTTT CATTTTTTTA CCTCTGGACT CTATTATAAT | 4440 |
| AAAAAATTGG AAGTTAGGAA ACTACGGAGC TAAAAAAGAA ATTAAAAAGA TTAAGCAAAC | 4500 |
| GCTTGACAA AATTTTAAAA AGTGCTATCA TAGACTATAG ATTATGAAAA TAATGAGGTA | 4560 |
| AACAGATGCA AGAAAAATGG TGGCACAATG CCGTAGTCTA TCAAGTCTAT CCAAAGAGTT | 4620 |
| TTATGGATAG TAATGGAGAT GGAGTTGGTG ATTTGCCAGG TATTACCAGT AAGTTGGACT | 4680 |
| ATCTAGCTAA GCTAGGAATC ACAGCAATTT GGCTTTCTCC CGTTTATGAC AGCCCTATGG | 4740 |
| ATGATAATGG CTATGATATT GCTGATTATC AAGCGATTGC GGCTATTTTT GGAACCATGG | 4800 |

680

| | |
|---|------|
| AGGACATGGA TCAGCTGATT GCAGAAGCTA AGAAGCGTGA CATTCGTATC ATCATGGACT | 4860 |
| TGGTGGTCAA TCATACCTCA GATGAACATG CTTGGTTTGT CGAAGCCTGT GAAAATACTG | 4920 |
| ACAGCCCTGA GCGAGACTAC TATATCTGGC GCGATGAACC CAATGACCTA GATTCTATCT | 4980 |
| TTAGTGGGTC TGCTTGGGAA TACGATGAAA AGTCAGGTCA ATACTATCTC CACTTTTTCa | 5040 |
| GCAAGAAACA GCCGGATCTC AACTGGGAAA ATGAAAAACT TCGCCAGAAA ATTTATGAGA | 5100 |
| TGATGAACTT CTGGATTGAT AAAGGTATTG GTGGTTTCCG TATGGATGTT ATTGACATGA | 5160 |
| TTGGCAAAAT TCCTGACGAG AAGGTAGTCA ATAATGGTCC TATGCTCCAT CCCTATCTCA | 5220 |
| AGGAAATGAA TCAGGCGACC TTTGGAGATA AGGATCTCTT GACAGTAGGG GAGACTTGGG | 5280 |
| GAGCAACTCC AGAGATTGCC AAGTTCTACT CTGATCCAAA GGGGCAAGAA TTGTCTATGG | 5340 |
| TCTTCCAGTT TGAACATATC GGTCTTCAGT ATCAGGAAGG TCAGCCTAAA TGGCACTATC | 5400 |
| AAAAAGAGCT GAATATCGCT AAGTTAAAAG AAATCTTCAA CAAATGGCAG ACAGAGTTAG | 5460 |
| GAGTTGAGGA CGGCTGGAAT TCCCTCTTCT GGAACAACCA TGACCTCCCT CGTATGTCT | 5520 |
| CAATCTGGGG AAATGACCAA GAATACCGCG AAAAATCTGC CAAAGCCTTT GCAATCTTAC | 5580 |
| TTCATCTCAT GAGAGGAACT CCTTATATCT ACCAAGGTGA GGAGATTGGG ATGACCAACT | 5640 |
| ATCCGTTTGA AACACTGGAT CAAGTAGAAG ATATTGAATC TCTCAACTAT GCGCGTGAGG | 5700 |
| CTCTTGAAAA AGGTGTTCCG ATTGAAGAAA TCATGGACAG TATCCGTGTT ATTGGACGTG | 5760 |
| ACAATGCCCG TACCCCTATG CAATGGGACG AGAGCAAAAA CGCTGGTTTC TCAACAGGTC | 5820 |
| AACCTTGGTT GCGGGTTAAT CCAAATTACG AGATGATCAA TGTCCAAGAA GCGCTGGCAA | 5880 |
| ATCCAGATTC TATTTTCTAT ACCTATCAGA AACTGGTCCA AATTGCAAG GAGAATAGCT | 5940 |
| GGCTAGTTCG AGCTGACTTT GAATTGCTTG ATACGGCTGA TAAGGTCTTT GCTTATATAC | 6000 |
| GTAAGGATGG CGACCGTCGC TTCTAGTTG TGGCTAACTT GTCCAATGAA GAGCAAGACT | 6060 |
| TGACAGTAGA AGGAAAAGTC AAATCTGTCT TGATTGAAAA CACTGCGGCT AAAGAAGTAC | 6120 |
| TTGAAAAACA GGTCTTGGCT CCATGGGATG CTTTCTGTGT GGAATTACTA TAAATATTTT | 6180 |
| TTGCAGAAAA ATTTAAAAAT GAAATCGTAT AAAACAAGG GAGGACTGTA TAAAAGACAG | 6240 |
| AAATCCTTTG TTTTATTATA CCAAAGTTTA TAACTTTCA TTCTTGAAAT TCAATTAACT | 6300 |
| TTACAAATTC CCACTATTAA GGAGAAAGAA GATGAACATA AAGAAGCGTG TCCTTAGTGC | 6360 |
| AGGCCTGACT TTTGCATCTG CTTTGCTTTT ACCCAAATCA TTCATACCTC TCTCAACTAG | 6420 |
| ATGTAACCTA CAAAACCCCT GACCTCATGA GCCACTTTCT TCCTCCTCAT GAGGTCAGTT | 6480 |
| TTACTTTCTG CTGTTCCAGT ATCGTTTTTC CTCGCTAGAT TTCTCAGAAA GGGCAGACTC | 6540 |
| CTCCCTTGGT GCGTCACACG ATTTTTTCAT CTCGACTGTT CTTTAATGCA TCATTAACGA | 6600 |

681

| | |
|--|------|
| CGCTTTTCTT CTAGGTGGTT CATAAGGAAC AGGAAGATTC AGGTTGACTT TTCTAATCCT | 6660 |
| AGAAATAAAGT GCTGAAAACA ATTCGGAATA GGCATAGAGA CTAGACAATT TGAGGAGCTG | 6720 |
| CTTGCGTCCT GTTCGAACAC ATTTTCCGG | 6749 |

(2) INFORMATION FOR SEQ ID NO: 85:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1842 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 85:

| | |
|---|------|
| TCTACCCATG GACTTTGAGG CATTCAATGT TCCATCTTCT AGTGGCGAAT CTTTGTATAC | 60 |
| AAACGATTCA ATTCATTGG ATAGTGAAAC TCTCCCGCAA ACATTTTCTT GGTAACTCA | 120 |
| ATCCAGCTGA TATTTCTTTC AGCCAAAATA ATGGACAAGT TCTCCCAAAA TCGTTCAGCC | 180 |
| ATATTGCTTC TCCTTTAGTT AGATAAATAA TGTGTTTGCG CCATGTAAAT CAATTGTTTC | 240 |
| GTATCTCTTG GCAATAGAGC TCTAGCCTCT TCCAAATTCA GACTTGGATA AACTCGCTTA | 300 |
| TTTGAAACCG CAAGAGGAAG TCTGATGGTT AGTTCAGGAT TTTTAAAAT TATCTCAACG | 360 |
| AAATCCGTTA ATCTTAGATT GTCACGGTTC TTAAATCGTA ATAAATGGG AGATAAAAC | 420 |
| TCAAAACAAT CTGAAGAATA GCTCATCATC TCAATTAATT TGTCCTTTGT CATTTAGAA | 480 |
| ACTGAATGAC AAGATACCTC TATGCCATAG TTTTGAAGA AATCTAAAAG AAGTTGATTT | 540 |
| CTTTGTCTAT TTTTACTTAG ATAGAGATCA ATCATGGGAG ACCTCCCAA GATTCGGTTC | 600 |
| CATTTGATAT TCTGACACGA TTAAGGAATC TAATAAATTA AGGAATCTAA TAAATTTGCG | 660 |
| AAGTTAATCG GTTCTTGTC TTCATCATAA GCTTTTACAG TTAATTGGGT TGTAAGTATT | 720 |
| CCCTCTTTTC CCTCGGCTCG ATAGCCTTGT CCATATAAAA CAAAACGAG ATTTTGATGA | 780 |
| TCATCTACAA AGGCATCAAC CCCATTCTTT ATGTCTTGAC TTTCAAGGAA TTCCATAACG | 840 |
| TTTTGAAGAT AGGATTCGTA AAATAGTGGG TAGTTATGTT TTTTATGGTA ATCATCTAAA | 900 |
| AATGTCACTT CAAACTCACA TGGAGAGTAA TTTTGACTTT GAACAGCCTA AAAGTGCCAT | 960 |
| CAAATTTGAA TTGAATAAA TCAAATAAAT AGCCCCATCC TCATCAATCC AACCTTTGCT | 1020 |
| CAAAGACAAC TCCAACCGAT CTTTAAAAC TGAGTAAACC ACCTTAACCT CCAGTTTCAT | 1080 |
| ATTCTTATAC CGTTCCTCT CAAATAAAG TTTGGGGAGC TTATAATAAC GCTCTGATGT | 1140 |
| CTGATATTGA TTAGCGGTAA TACGCTTCAT TATTGTCCCT CCAAGACTAA AATTCCAACA | 1200 |

682

| | |
|---|------|
| TTTCCAAATT CATCAAATCG GATTAAACCT ACTTGTTCCTCA TTTCATCAAC TAACTGAGTT | 1260 |
| GCTTTTACCC AAATCATTC A TACCTCTCTC AACTAGATGT AACTTACAAA ACCCCTGACC | 1320 |
| TCATGAGCCA CTTTCTTCCT CCTCATGAGG TCAGTTTAC TTTCTGCTGT TCCAGTATCG | 1380 |
| TTTTTCCTCG CTAGATTTC TCAAAAGGGC AGACTCCTCC CTTGGTGCGT CACACGATTT | 1440 |
| TTTCATCTCG ACTGTTCTTT AATGCATCAT TAACGACGCT TTTCTTCTAG GTGGTTCATA | 1500 |
| AGGAACAGGA AGATTCAGGT TGACTTTTCT AATCCTAGAA TAAAGTGCTG AAAACAATTC | 1560 |
| GGAATAGGCA TAGAGACTAG ACAATTGAG GAGCTGCTTG CGTCTGTTC GAACACATTT | 1620 |
| TCCCACCACG TGAAGAAAA GATGGCGGAA GCGTTTGATT GTTAAAGTTT GGAAGTCACC | 1680 |
| TCCAGCTAGA TGTTTGAGAA AAAGATAGAG ATTGTAGGCG ATACAGCTCA TCATCATACG | 1740 |
| AATTCGTTT TTGATTAAGG TTGAACTATC CGTTTATCG CCAAAAAATC CCTCCTTCAT | 1800 |
| CTCCTTGATG AAATTCTCGG CTTGACCACG TCCACGATAA AG | 1842 |

(2) INFORMATION FOR SEQ ID NO: 86:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 19390 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 86:

| | |
|---|-----|
| TCATCTTTAT CTCCTCGAAA TTTTCTAATA TAGCCATTAT AACAGAATTT TGTGAAAATT | 60 |
| CCTATTATAG TAAATCACTA TTTCAGTATA AAAAGAAAAA ACGAATCAGA CGATTGCTC | 120 |
| TTCTTAAAT CTGAAAATAG CTTTCCAGAA AGGATTAGCC GATTTTTTGC AGATTGAGCA | 180 |
| CTGCATCGTG ACTCATCAAG ACTTGACCAT ACTCTTGTA GACTGAGCGA CTGATATCAC | 240 |
| TATCGTCTGC AAACGCGC ATACGGGCA ACAGCCAAGC TGGATATGGG CTTGGATGAT | 300 |
| TTTCAATATC CACTAAAATG GTCAAATAAT AGCGCTCGTT CATTTTGTAG AGTTCAGAAG | 360 |
| TTTCCATTTT AAAAGTCACT GTCTTGCAA AAGTACCAA GTCAGCCAAC TTAGCAAAAG | 420 |
| AAAGGATGTA GTAGATGTAA GGTCTTTCT TACTCTCAGC TTCTTGTTCA GCCTGCTCTT | 480 |
| GCTCTTCTTC CTTGACTTCA ACTTGCTCAA GAGATTGAAT GGCTTCGATA TCATCCTTGG | 540 |
| TTTTGTCTGC GATGCTTTTT TCCAGGGTTT TGATAAATTC ATCTGGAGAC ATTTGAGCCA | 600 |
| ATCTTCCAT ATCTGGCAAA TCCGATAAGT CTTCAAAATC TAGATTTTGG TCAATCTTTG | 660 |
| ACTTGTCAC AAAGACATCT ACCTTATCAG GTTTTGGAGT CACACGGAAG CTCAACATGC | 720 |
| CTGTATCCAG AAAGCTATCA GGCATCTCTA GTCATCCAA GATAGCATAA AAGAACTCTT | 780 |

683

| | |
|---|------|
| CTGTTTTTTC TTGAGGAACG AGAAAGTCAG CAATCTCCAT TCCACGATCC ATCAAATCCT | 840 |
| CTAAAGATAT CGTGATTTT AAAGTTGTAT CACTAATTG TTTCATTTTC ATTGCTAGTA | 900 |
| ACCTCATACT TTCAGTTCTA TCTATTATAC TAGATTTTTA CGATTTTATC AAAAGAAGGC | 960 |
| TCCTCTATAC GGATAGATT TCCCTAGGGT CTTTCTATAG GAGACTCCAA AAGAAAATT | 1020 |
| CTGCAGACAG ATAGAAAAAG CCTTCAAAAT CGGCTAAGAG CCGACTTTGA AGACCTTATA | 1080 |
| CATCAGAATA CTTATAATTT AAAGGTTGCT ACACCGAGGA TAGAACGATT TAAGTTTCTG | 1140 |
| AGAATTGAA GACTTTGCTC AAATTTCTTA TAACGAGTCA CTCCGTACTC TTCAACAAGA | 1200 |
| AGGACTGTAT CTCTTTCCAA AAGAGATGAT ACATCCTGTA AATCTACAAA ATGCATTCTT | 1260 |
| TTTAAAGCTT CTTGACTCTG TTTCAATTTA TCTAAGATAG CTTTATTTGA GCTAACGATG | 1320 |
| GTCAATTCTT GTCCAGTATT TTTGTATGAC AAAACATCTG CTAGGTTAGC AATGTGTGTA | 1380 |
| ATCTCTGTTA CAAAATCAAT TTGATACTGA GAAAAATCAC CTACTCTATT GATTGTGGA | 1440 |
| TTAAAGAGAT AAATAACAC ATTTCCCATC ACAACCAAAA TCACACAAAC CACTCCAATA | 1500 |
| ACAATAAAC GAAGAATCAG ATTTTTCACA TTTAAGCCAA GCGCTGTTTC ACCATTTGCG | 1560 |
| TTCAATTCTT TAGAGTTGAT GGTTCAGT TTTTCAATTT TCACATTTGC ATAGGCATGT | 1620 |
| TTAAATTTCT CAATCAACCC ATCAATTTTT TTCTCTAACA AGTTATTGGC ATCTTTACTT | 1680 |
| GATGTCAAAA TTTTCACACC AACCCTGCA TCGTCAATCA TATAGTAGAC GGTCAATTTT | 1740 |
| TTCCACCAAT AGTCATTCTG TGAATTTTC AAGGTTGTTT CTGTCGTGTC TAATTCAGT | 1800 |
| GCAATTTTTT TCAACTCACT GGGTTCTACA TCATTGAAA GATAAGCTCC ATTCAAATTA | 1860 |
| CCATCAATCA ATTTCCCAT AATACTACTA TAACCACCA TTTGATGATT CAAAATCGTT | 1920 |
| TTGTCCGACT CTTTGGAGG AGTGATTTTA TAGATAAGAT AAGTTGAATA ACTTGTGTA | 1980 |
| TCTTTGACAG TGTTTTTATT CCTAACTGCT TTAATTGTAA ATGGTACAGC AATGAGAGCA | 2040 |
| AATAAAGCGA TGAGAGCTAA AATATTTGCT TTTCGCTTTT TATAAAGATT TGCAAACAAA | 2100 |
| TCAGTACTG AATAATGTTT AAACATGATT TTTTCTCCT TTGTTTAGTA GATACTAGTT | 2160 |
| TTCTTTGTA AGCATTTTTG CTACAAATAT AATCACAAGA ACAATTCCCC AGAATTGCAT | 2220 |
| TGTAAATAAA TTGAAGAAAC TTTCTGAAAA GCTGCTTCTT GGCATAAAGA ATAGATTATT | 2280 |
| CAAGATGAGT AGGGATAAAG CAAATAGGAT TGTCCTTGAG CGATAGGCTA CTTGCAGCAT | 2340 |
| GGCTATAAAT AATACGCCGA GTAAGAACT AAGCAGAAAG ACTCCAATCA TACCATAGTC | 2400 |
| GGTATACAAC TCCATGATAT AACTACTTCC GATACCATGC CCTTTCAAGT ATTCCTTGTT | 2460 |
| CAAGACAAGA TAGGATAGAT TGTGGGCATA ACTATTACTA TCAATAGCTA GTTCCACACT | 2520 |

684

| | |
|---|------|
| ATTGGTTGTA TGTTCAAAGG CTTTTCCTCC GAAAATGGCT CCCAACTCC CCCTTGCAAA | 2580 |
| ATAATCAAGA ACAGGACCAA AAGTAAAATT ACGGAAATCT CGGTAAGGGA GGCTACTGTT | 2640 |
| AAATAGAAAA CCTCGAGCCA GAACACCAAA ACTAGTCCCT TGTTTATAGA TAAAGTCAAG | 2700 |
| TAAGATATCC CAGAAACCTG TATGGGAAAC TTGGACATTA TCCCGTACAT AATTGAGTAC | 2760 |
| TCCCATCGCT AACATGAGAA TAGGAGAACC TACAAAAATC GCTAACTTTT CTTTAAACCC | 2820 |
| AATCCATTTT CCTTTTTCAG TTTGCTCCCG CATAAAGTAA TAAACAAAAG CAAATmAAAT | 2880 |
| ACTTAAATAA AAGGGATTTC GTGTCCCAAT TGCCAAATGA ATAGTATTAG CTGCAATAAA | 2940 |
| GGAGACAAGC ACTGCTGTGG CCTGCAATTT CTTGGCTTG GTTGCCAGAT ACATACACAT | 3000 |
| TGCATAGACC GTAAAGGTAG ACAAATGTA GGTAAATAA GGCAGTTTAC TTTCAAAATT | 3060 |
| TGCATAGTAG GCATAGTAGG AAGTCTGCAA ACGATACAAG AGCCGTTCAA ATAACCGAAT | 3120 |
| GAAATAGAAA GGATAAGTTA GAAGAAAAC TCCTAGTGAT ACAAGCGTA ACCGCTTGAT | 3180 |
| ATAAACCTCT TTTAGAGAAT TTCCTATATT TGCTACTTTT ATTTCTTCC TAGCTATGAA | 3240 |
| GTAACGAGCC AGAATGCCTC CTGTGGTCAA GCCCAGAATC GAAATCATGA CAACTATAAA | 3300 |
| GGCAAAACGA TAGGCTATTG GATGATAGGT ATCCAAAGCA CCATCCCTAA AATAATCAAT | 3360 |
| GGTCGGTCTT GATACCAGAA ATACAAAAAT GGTAAATAG AAAATAAAAT GGATTAAAGTA | 3420 |
| ATACTTGATA TCATTCCAAC AAGCAATTAA GCTACTAACC AACAAGAACA ATAAAGTAGA | 3480 |
| AAGTAAGCTA ACATTATTAT TATTAAACAG ATACACAATT CCACCTACTA GCGTCAAGGC | 3540 |
| ATAACTGACT ATGGTCAAAC TAAATAATAA TCGTTTCCCA TCAATCACTT GGTCACCCCC | 3600 |
| GTTCTAATGT AATTTTTCAG ATTTTTCAT ATTTTTCAGT AATAAGAATC GATATAAGGA | 3660 |
| AATATTTATG AATAGGGCCA AAGCACTAAT TCTTCTCCCC TTACGGAAAA TTGGATTCTT | 3720 |
| AGAAATAGCA AAGGCATGGC CTTTAAAAA ACGATGAATC TGAGAATAGG CTTCAAACCTG | 3780 |
| TTTATACTGA TCATCTAGCA ACATCTTATC CAGAATAAAG AAGTGGGCAT AGGCCAATCT | 3840 |
| GAAAAAAGCG ACCTCTTTCA AGTCAGGATA GTTTTTCACA ACTTCATTAT AAAACTTTTG | 3900 |
| GTAGATATCA ATATAGGCTA AATCCTTCTC TGCATAGGGT TTGGTCGTAA TACTATCCCC | 3960 |
| TCTATGGAAA TAGTAATAAT AGGGTTTAGT ATTAACCACA TACTTCTTGG CCAACTTGAT | 4020 |
| TAAATCAAAA TGGTAATAGG CATCTTCGTA AATCAACCCC TTAGGAAAGG ATAGGGCAGT | 4080 |
| TGCAATCTGT CTCTTGATTA GCTTATTGCA AATCGTCCCA GGTATTTTTT CACCTATGAG | 4140 |
| GTATTCCTTT AGAAATGTTT GAGAATCACA GACAAAATAG TCATCCTGAT TGGCTGACTG | 4200 |
| TGGGCTTTCA TCATTAGCAT AGACATTCAT GACACCACAG CTCGAAACAT CCGCATCTTC | 4260 |
| TTGAACTAAT TGCTCATATA AGCTCTGAAT CATTTCTGGA TGGATATAAT CATCTGAGTC | 4320 |

685

| | |
|--|------|
| AATAAAATC AGATAATCCC CGTGAGCCTG CTTCATCCCA TCATTTCGTG CTTGCGACAA | 4380 |
| TCCTTCGTTC TTTTATGAA GCACTGACAC CCTGTCATCT TGTTCAGCGA TTGAATCACA | 4440 |
| CAAGCGACCA CTTCATCTG TTGCACCATC ATCAACAAGA ATAATTTCCA GATTTTGATA | 4500 |
| GGTCTGCTTC TGAATGGAAG CTATCGATT TCTAGGTAC TGCGCCACAT TATAGACTGG | 4560 |
| CACAATCACA CTAATTAATG CAGTTTCCAT GCTACTCCTC TAATAGTTTT TCTACTTGTT | 4620 |
| CGATTTGTTT TGTAATTGTA AATTGTTGAA TGAATTGGCT AGCCTCATCG ACATCAAAGT | 4680 |
| TTGAGGCAGA AGTCATGTAA TTAGTAATCG CCTGAGCTGC CTCCTGATTG CTCTCAATGA | 4740 |
| TTTGTCACAA TCGTCCCTCT TGGGATAATT CCTCAGCCCC TCCAACGTCC GTAGAGATAA | 4800 |
| AAGGGAGTCC CAGACTCAAG GCCTCCACAT ACACTCCAGG AAAACCTTCT TGTTTAGACA | 4860 |
| TAGACAAAAG AACTTTCGTC TGAGATAGAT ACTGATAAGG ATTTTTTTGA TAACCAAGGA | 4920 |
| AATGTACATA GTCCCTCAATC CCATACTCTT TGA CTGCTTTT TTTCAAGTCC TCTTCCATAT | 4980 |
| CACCAGCCCC GATAAAATAG AGATGATAGT TTTTCCCTC TTGGTGTAAT AATCGTATCA | 5040 |
| CTTCCACTAC ACGGTCAGAA CCCTTATTTT CCTCAATCCG TCCGATAGTA CAGATACTTT | 5100 |
| GAGGAGCAAT CTCGATATCG ATCTTCTCTT GAGATTTTTC TAGAATAGTC TGAAAATCAT | 5160 |
| ATCCATTGTA GATTGTCTGT AATTAGAAG TATAATCTGG ATAAACTTCC TTGATAGAAT | 5220 |
| TGCTGGTCTT TTTTGAAATC CCTACAATTG TATTGCGAGC ATCCAAGTGG CTTCTATGTG | 5280 |
| ATTCTCTTTT AGAGCTATCC TTAAGAAGT CTTCAATACT TCCATGAATC CAAGATATCT | 5340 |
| TCTTGACTTC TCTTCTTTTA GAGAACAACA GTGGTGATT CATAATGGTA AAAGAACTT | 5400 |
| CAACATCATA ATCATCTTTT ACAAGCAAAC GACGAGTCAG TCTTGAAAAA TAAATTTCTCA | 5460 |
| TTCTCCACAA AAAAGCTCGT AACCATCTGG TTTGGCGATA ATCTTGAAGG GATTTTAAAA | 5520 |
| TGCGTACATG CTTTGGAACA GATTCATATC CCTTGTCAA GTGCTCCATT TCAAGAATAT | 5580 |
| CAATATCATA CTTTCTGGA TCCAGATTTG AAACAATGGT TGATAGAATC TTCTCTGCAC | 5640 |
| CACCTCCAAG AGAAAAAGAC CACATAAAAA ATAAGATTTT TTTCTTAGCC ACCATATTCT | 5700 |
| CCCTTGATTT CTGTATAAGA CTTATCCATA TCAGCGATGA CAGCATCATG ATGCGGTACC | 5760 |
| TGCTTGCTCG CTGGTGGAGG CGTCATATAA TCCCAAAAAG CAGTTCTGAG ATAGACATCA | 5820 |
| TAGCCGATTG GAATAGGCAT CTCTGTTCTT TCAAATGGCA AGAAAAGATT GTCTTCAAAA | 5880 |
| GATGTGATTG GGTACTTGTT TCTCATGTAG CCAGGACCTG AGCATAATTC TGTAAATGCCA | 5940 |
| TCACAATCAG CCAAATCATA CTTAGTCATT TCTTCTCAG CTTTTTTCCA GATGCGATAA | 6000 |
| CGGAGAGATT TTGGAGTCAA ACCCAGTAAA ATGCGACTTC CCCATTTTCAT GAGATCACCA | 6060 |

686

| | |
|--|------|
| TGCTTTTCTG GAATAGTTTG CGCACAAAAG AGTGAATAAA TCAAGGCCCA ACGAACCTGT | 6120 |
| TTTTCCTCGT CAGCTGGATT TTTTCGGATAA TAATCCAAAG GCAAAACATC CAAGGCCAGA | 6180 |
| CCATGTGGCA AATCCAAATC CTGCTGATAA GGCTTGATAC AGGTGGTTTT CTTGTCACGA | 6240 |
| ATGGTAATAA AAAGATTACG ATCAACAAA TCCTTGTGAC TCTTTGACAA GAAATAACGT | 6300 |
| TCATCTGCAT AACGAGGCCA TAATCTGCT AATTTCTCAT AATCTTTACG AGGCATAAAA | 6360 |
| AAGTCTAGGT CGTCGTCCCA AGGAATAAAT CCCTTGTTC GAAGGGCACC AATAGCGCCT | 6420 |
| CCGCCACAGA GATAACAGAG CAAATCATGT TCTTTACAAA AGGCCACAAA ATATTCAGCC | 6480 |
| ATCTCCAGAC TACGAGCCTG AATTGCTTTT AAATCAGTCA TATTGTTTAT TATTCTTTCT | 6540 |
| ATCGTATCGT TTCATTATAC CACAAACAAG GGGTGAATAT CTATTGCAGA CTGTAAAAA | 6600 |
| TCAAAGCCTG ACTGCTATCC AAATAGCTAT CAAACTTTGA TTTTCTGTC TTATACTCTT | 6660 |
| CGAAAATCTC TTCAAACCAC GTCAGCTTCA CCTTGCCGTA GGTATAGGTA ACTGACTTCG | 6720 |
| TCAGTCTTAT CTACAACCTC AAAACTGTGT TTTTAGCAGC CTGCGGCTAG CTTCTAGTT | 6780 |
| TGCACTTTGA TTTTCATTGA GTATTATCTT ATCTTAAGCC CATTTGAGCG AGCTTGGTTT | 6840 |
| GATATTTGTT TTGATCAACC AGCAGGCCCA AGCCCCATA AACATCATAG GCATCTACCC | 6900 |
| AGTCACCCAG TTCTGGAATC GTCAATTTTT CAATACCATT TTTTGCTCCA TCCAAAACAG | 6960 |
| ATAAACCGTT TGTTAGGAGG AAAGTATAGG GTACGTTGGT TGAGGTCATA GCAAAAACCT | 7020 |
| TTCCAAGAGC TTCAGAACCA GTGAAAAGTT TAGTGGGATC TTTAATTTGC TCTAAAATTG | 7080 |
| CTGTTAAAAC TTGTGCTGT CTTTTGTAC GGCCGTAATC TGCCTCATCA TCATCACGGA | 7140 |
| AACGAGCATA ATTGAGCAGG GTCGAGCCAT TCATCTGCTG TTTTCCGACT TTAATGGTTT | 7200 |
| GGGTGGAGA CTCAGTCTCG GTAGCGTATA AATCATCTCC GACTGTAGCT TCTGTTAGGG | 7260 |
| GACGCCCATT CAATGTTGAA AATTGAGCAT CAATCGTCAC CCCATCAGGG AAAAGCGTGT | 7320 |
| CAATCGCTGT GGCAAAGGCC TGGAAATCAA CCAAGGCGTA GTACTTAATG TCCAAGTCAA | 7380 |
| AATTATCTTT CAAGACTTGG CGAACCATT CTGCCCCCTT TTGCCCCCTT TGTTCCTTA | 7440 |
| ACTCGTAGGC TACGTTTAACT TTGTTATCTG TCTGTTTCTT ACCATTAATC ACTTGACTAT | 7500 |
| AACCATCTAT ATAGACCAAA TTATCACGCA TGAAACTGAC TAGCTTCATT TTCTTATCTG | 7560 |
| AGCCCCGAC ATTTAATACC ATAATAGAGT CAGTTCGTGT CTCAACACTG TTCTGGCCGA | 7620 |
| TTTGACCATC AGTACCCATG ATTAATAATAT TAACTCCATC TCTAGTGTCC TGACCATTAA | 7680 |
| AGACTTCTAC TTGAGCTGCC CGGGCATCAG CAGTTTCTT TCGCTAGCA TCTTGTAAC | 7740 |
| CACGCAAAAA CATGAATACC ATGGCCAAAG CCACACAGAC CAAAAGTGAA AAAATCACCA | 7800 |
| TAAAAATTCG TTTAAGACGG AGCTTCCGTC TTTCTTTTT TGGAGGGAAA GAGAGTGCTT | 7860 |

687

| | |
|---|------|
| GTGATTTGGA TTGTGAGCGA CTCCGGTTTC CATAGCTTGG TAAGTCAACC TGCTCTTCTC | 7920 |
| TTTCTTGTTT CAAGCTAGAG CTACTATTTC CCCTAGCAAG AGTTAGCTTT TCTTGCAAAT | 7980 |
| AGGCAAATC ATTTTTTTCT CTCTCATTGA GATAGTGAAT ATTTTTTAGC AAATAATCAT | 8040 |
| AACGCAACTG CTCATGATGA CTTAAGGGAT TTTCTTTACT CATCTTCTCT CCTTTCCATG | 8100 |
| GTCTGATATT GGATAAATAG GATAGGCACC CAGAATTTTA TACTGGATTC CAATCGCTTC | 8160 |
| TAATTCCTTT TGGGCAAAGT GGACCAAGTC CTTATCGGTA TAATCCACAT CGATAATGAA | 8220 |
| AAAGTATTCA CCCAGTGCTG TCTTGAGTGG ACAACTTTCA ATTTTGTGCA AGTCAATTCC | 8280 |
| TCGCCAAGCA AAGGTCGACA GGGCCTTATA AAGTGCACCT GGAAGGTTGT CAGGTAATGT | 8340 |
| CAAGGCCAAA CTCATCTTTT CAGTTTGTGC TTGCAAGGGA ATACTAGGCT TTTCAGCTCC | 8400 |
| TAGAACCAG AAACGTGTGA AATTGGCTTC CATTTCTGTA ATATCCTCGG CAATCAGTTC | 8460 |
| CAATCCATAT TCTTCAGCAG AACTTCTAGG TGCAACTGCT GCAAAGGGCT GGTCTGGATG | 8520 |
| TTTCGAAATA AAACGGGCCG CATAAGCTGT ACTAGCTGTT ACCTCGATTT GAGCCTCTGG | 8580 |
| ATATTGTTCA TCGATGAATT TCTTTCCTTG AGCCAAGGCC TGTGGATGTG AAAAAATCTT | 8640 |
| TTCAATCTTA GTATGGCCTG GAACCACCAT CAACTGCTGA TGAATAGGCT GAACGATTTT | 8700 |
| TGCTACTGCT TGGATGTGAG CCTGATGAAA AAGATAGTCC AAGGTTTCAT GAACACTACC | 8760 |
| CTCAATAGAA TTTTCAACTG GCACCACAGA ATAGTTCACT AATCCTTGCT CATAAGCCTT | 8820 |
| GATGACATCT GTAATGTTGG CAAAAGCCTG CAATTCCTCA TGAGGAAAAG CTGTCTGCAC | 8880 |
| AACGTGGTGT GAAAATGATC CCTTGGGACC TAGATAAGCA ATTTTCATCT TAGTTCCTCT | 8940 |
| ATAATTTCTT CTGGGCTTAG CTTGGTCACA TCCAAAACCC GACTAGCCAC TTCCTCATAC | 9000 |
| CAAGCCTGTC TTTCTTGAA AATAGCTACT AGTCTTCTCT TGCTATTATT TAGAAAAAGC | 9060 |
| GGTCGCTGAT TGTCTTATC AGCTGCGATA CGTTGGTAGA GGGTTTCAAA ATCTGCTCTC | 9120 |
| AGGTAGATGT TATCTGTATT AGTCTTGAGT AAGTCACGAT TTCTCTGAGA AATAACCACT | 9180 |
| CCTCTCCAG TTGACACGAC TTGGTCTGTT TGTAGTAAAT CAGCTAGGAC TTCTGATTCT | 9240 |
| ACCTGACGAA AGGCTGTTTC TCCCTTTTCA GCGAAAAAT TCGCAATGGA CATACCTAGG | 9300 |
| CGATTCTCAA TCAGAGCATC CATATCAAGG TAATTAGGGT CCAAGCCTCT TGCAATAGTC | 9360 |
| GATTTTCCAG CCCCCATAAA CCCTAATAAC ACCTTAGCCA TGAATCAAGC TCTCCAAATC | 9420 |
| ATCAAAGAAA CTAGGATAGC TGGTATTGAT GGCTTCTGCA CGGTCAAGCT CCACCTCTCC | 9480 |
| ATCTGCAACC AAGAGGGCTG CGATAGCTGT CATCATGCCG ATACGGTGGT CACCAAACGT | 9540 |
| ATTGACTCTA GCACCGTAA GAGCTGATTT TCCTTTGATA ATCATCCCAT CTGCCGTAGG | 9600 |

688
AGTAATATCT GCTCCCATAC TATTTAAGGC GTCTGCCACA ACCTGAATAC GGTCTGTTTC 9660
CTTGACCTTG AGCTCCTCAG CATCCTTGAT AACTGTTACA CCTTGGGCTT GGGTCGCAAG 9720
CAGGGCAATA ATGGGCAATT CATCAATCAA TCGTGGAATC AAAGCGCCAC CAATCTCTGT 9780
TCCTTTCAAG TCAGAAGACT CAACAATCAA GGTAGCAGAT TTAGCGACTG GATCGATTTC 9840
AGTTATTTCC AATTTTCCAC CCATGGCAGC AATGACATCA ATAATACCGG TGCAGATTTC 9900
GTTGATCCCC ACATTCTGCA GCACTAGACG AGAATTGGA GCAATCAAAC CTGCGACTAA 9960
CCAAAAGGCT GCACTGGAAA TATCTCCTGG TACGACCACC TTCTGTCTTG TCAATTTTTG 10020
TGGCCCCCTGG ACTGTGATTT TCTTACCATC CACACTTAAA TGACCACCAA ATTGTTTCAA 10080
CATATCTTCA GTATGATTAC GGGTGACTC TTTTTCGATA ATAACCTGACT CCCCTTAGC 10140
TTGTAAGGCT GCAAACATCA AGGCTGACTT GACTTGGGCA GAGGCAATTG GCAACTCATA 10200
ATGAATAGGT CTTAGGTTTT TCGTCCCTTT TAAGCGAAGG GGAGGCAAGT CTCGTTCACT 10260
TTGCCCTGAA ATGCTGACGC CCATTTTTTT CAGTGGGAAGG GTCACACGGT CCATAGGACG 10320
TTTGGAAGA CTATCATCTC CAAACATCTC TACTTCGAAA TCTGCACCAG CAAGGACACC 10380
TGAAATCAGG CGAATCGAGG TGCCAGAATT TCCCATATTA AGGGCATTTC GTGGCGCTTT 10440
TAAGCCAGCC ATGCCTACAC CTTGAATGGT AATAACCCCA TCTTTATCCT CAATTTCAAC 10500
ACCAAGGTCA CGAAAAACCT GCATGGTCGA AAGAACGTCT TCACCTCGCA GAATATCATA 10560
AACCTTGCTC TCACCCTCAG CCAAACCTCC AAAGATAATG GAACGGTGGC TGATAGACTT 10620
GTCACCTGGG ACGCGGATAC TACCATGTAA ATGGCGAATG TTTGTTTTTA GTTTCATACT 10680
GGACCTCATA CTTGCAATAC TTTTACCTAT TTTATCATAA AAAGCCAGAA ATTCCTTAAA 10740
AATTCCTGAC TTTAGGATCG TTCTTTTCTT ATTTTCAGCA TTCTGAACT GGTTCAAAAA 10800
CAATTTTTTC AATATCAGAA AGGTAAATGG CCAATTGTTG TTGCTTGGA AAGAATCTG 10860
ACAAGAGGCT ATTCCTTGA ATCTGTTTAC CAAAGCCTTC CATCTTAGCT TGAAGGACG 10920
CATCTGGCAT TTGACCTGTC TGTGCTAGT TTTGAATTC CTCTTGAAAG GCAAGATAAT 10980
CTGTAAAGAT TTTGCTTGCC TCAGCATCTG CTGCAATCGC ATCTTTAGCT GCTTTAACAG 11040
CCTTGATTC TGGTAATCCG CGTAGACCGC GACTGAGTTC GTTGCACCTA TCGTAAATAT 11100
TTGACATGTT CTTCTCCTTA TTTGATGACG ACTGTATAGT CAGTATTTTC TGTTATGAGA 11160
TGCTCAGCTC TTTCCAAGTC TTGAGCATTT TTAATGAAA TTTGTAGGAT TCCGTGAATA 11220
TCCTCAGCAT TTTCTCTGTT GATGTGGATA TTAACCAAGG AAGTTCCACG TAGCAGTTCC 11280
AAAATCCGCA GGATGACATC TTCTTCATCA GGAACGTCAA CATAGAGGTC GTAAGAGCTA 11340
TCCACACCAC CACGCTTATG GATTTCCATG GTCTGGCGTT GTTCACGCGC TTGGTTAAAA 11400

689

| | |
|--|-------|
| AAGTTCCAAA TTTGCTCTTC ATCTCCCTTA CTAATGGCCT GACCAATCGC TTCCAAACGT | 11460 |
| TCCTTGAAAT CCTCAATTCT ATCCAGAATG ATCTCGCTAT TGGACAAGAG AATGGAGGTC | 11520 |
| CACATTCTTG GCTCGCTTTC CGCAATTTCG GTCATATCTC GAAAACCACC TGCCGCAAAG | 11580 |
| CGCCTTGCCA TCTCATGCTC TTGAGCATAG ACCGCAGTCT GCTCCATGAG ACTAGAAGCC | 11640 |
| AAAATATGAG GAAAATGGCT AATCTGAGAA GTGACACGAT CATGCTCCTT GGCATCAATC | 11700 |
| TCGATAAAAC GAGCATGAAG ACCTGAAAGC AGATCCTTCA TTTCTTAAG CGTGTCTGA | 11760 |
| CTTGTCAGGC TTGAAGGTGT AAAGATATAA TAGGCATTTT CAAAAAGATT GACATCTGCC | 11820 |
| GAAGCAGCCC CTGTCTTGTG ACTACCAGCC ATGGGATGGG CCCCACAAA GCGAACAGAC | 11880 |
| TTGCCAGCCA AATACTGCTC CGCCGCATCC ACAATGGTTG ACTTGGTCGA ACCAGCATCT | 11940 |
| GAAATAATAA CGCCTTCTCG CAAATCCAAA TTGGCCAACT CCTTAATGAA AGCAATAGTT | 12000 |
| TGTTTGATTG GCAAGCTGAG GATAATGACA TCTGCCAAAG GAGCAAACT AGCAAAATCA | 12060 |
| TCCGTTGCAC GGTCAATCAT ACCTTCTTTC AAGGCGATAT CTCTCGAAGC TTGACTACGA | 12120 |
| TTATAACCTA AAATTCATA ATCTGGATGA TCGCGTTTGA TACCAAGTGC CATAGAGGCT | 12180 |
| CCAATCAACC CAAGACCTGC GATATAGATT GTTTTTGCCA TAGGAACTCC TTAATAGTTC | 12240 |
| TTTGATAGT CTCGGTGTTC GGCTACCGCT TCTTTTAGTT CCTCAAGATT ATCTGATGAG | 12300 |
| AATTTTTCGA GGATTTCTTG CGCCAGAACC GTTGCTACAA CTGCTTCCAT GACCATTCTT | 12360 |
| GCAGCTGGAA GAGCAGTCGG ATCACTTCTC TCCACGGTTG CCTTGTAAGG TTCGTGGGTT | 12420 |
| TCGATATCCA CACTCATAAG AGGTTTATAA AGAGTAGGAA TGGGTTTCAT GACCCACGA | 12480 |
| ACAACGATGG GTTGCCCATT AGTCATACCA CCTTCAAAC CACCTAGATT ATTGGTACGG | 12540 |
| CGAGTATAAC CGTCTTCTTT AGACCAGAGA ATTTTCATCCA TAACTTGGCT GCCTTTACGA | 12600 |
| TAACCAGCCT CAAAGCCAAG ACCAAATCC ACCCCTTTAA AGGCATTGAT AGAGACAACA | 12660 |
| GCTTGAGCCA ATCTTGCAAT CAATTTTCTA TCCCATTTGA CATAGGAACC AAGACCAACT | 12720 |
| GGAACGCCTC CGACGACTGT CTCCACAACC CCACCGATGG TATCACCATC ACGTTTGATT | 12780 |
| TGGTCAATAT AGTCCTTGAT TTCCTGTTCT CGTTCTTGGT TGACAATAGA AACTTCAGAC | 12840 |
| TGGGCAGCTC TTTGCTTAAT TTCAGCGACT GTCAGATTTT CAGGAACATC GATTTCTTGT | 12900 |
| CCACCAAAGA CCACGACATG GTTGGCAATC TCCATATCCA GCTCAGCCAA GAGGCGTTTG | 12960 |
| GCTACTGCAC CAACTGCCAC CCGCATGGTG GTTTCACGAG CTGATGAACG CTCCAAAGAA | 13020 |
| TTTCGCAAAAT CATCAAAACG GTACTTAATC CCCCACCA AATCGGCATG ACCTGGGCGA | 13080 |
| GGATGAGTAA TTTTCCGCTT GCTTTTAAGG CGGTCTTCAA TGTCTCCGC AGACATGATG | 13140 |

690

| | |
|--|-------|
| TCCAGCCATT TCTGGTGGTC CTTATTGATG ACATCCATAG TAATAGGCGC CCCTGTCGTC | 13200 |
| TTCCCGTGGC GAACGCCCGA AGTAAAGACA ACCTGGTCAT TCTCAATCTT CATACGACCA | 13260 |
| CCACGACCGT AGCCACCCTG ACGGCGTCTA AGGTCCTCAT TGATATCCTC AGCTGTCAAT | 13320 |
| GGAAGTCCAG CTGGAATTCC CTCAATAATA GCTGTTAGAC GGGGGCCGTG TGATTCTCCT | 13380 |
| GCAGTTAAAT ATCTCATACA CTCTCCTTAT TTTACCAAGT AGTCTTTCAT CTCTTCCAGA | 13440 |
| GAAACTGGGT GAATGGTCGC TGAACCAAGC TCTGGCACCA AGACCAATTT CAAGGTGTTA | 13500 |
| CCACGCGCTT TCTTGTCATG AGTAAGAGCC TGATAAAGCT TGCCAACTTC CCAATTTTCA | 13560 |
| TAGTCAACAG GCAAACCGAA TTTCTGACAC ATCTCTGTGA TAGATTGGGT AATGCCAGCT | 13620 |
| GGCATGAGGC CTTTTCCTC AGCAACCTTG GAAATCTGTA CCATTCCCAT GGCAACAGCC | 13680 |
| TCTCCATGCA TGACCTTGCC ATAACCGCA GTCGCTTCGA TGGCATGGCC AATAGTGTGG | 13740 |
| CCAAAATTGA GGTAAAGACG AATACCATTG TCCAACCTCAT CTTCAACCAC CATCTTGCGC | 13800 |
| TTCACCTGAC AAGAATGTTT AATCAAGGTC TCTGCATGTT CCAAATACT CTCAACAGAA | 13860 |
| CCATTGAGTC CCGTCAAGAG AGCCCACAGT TCTGGATCCT CAATCAAGCC ATACTTGATA | 13920 |
| ACTTCACCCA TCCCTTCAAT CAACTCTCTT TTTCCGAGGG TTTCAAGAAC AAGTGGATCA | 13980 |
| ATCAGAACCC CATCTGGTTG GGCAAAGGTC CCCACCATAT TTTTAGCAAA TGGTGTATTA | 14040 |
| ACGCCTGTCT TTCCACCGAT AGAAGAATCA ACCTGAGCTG TCAAACCTAGT CGGAATCTGA | 14100 |
| ACAAAGTGAA TACCCCGCAT ATAGGTAGAG GCTACAAATC CAGCCAGGTC CCCAACAACG | 14160 |
| CCACCACCAA GAGCAACGAT TCCATCGCTA CGAGTCAGAC CTTGCTTGAC TAGAAATTCA | 14220 |
| TAGACTTTCT GAACAGTAGT TAAATCTCTT CTTTCTTCAC CTTCTAAGAA ATCAAAAACA | 14280 |
| GCTACCTGAA AACCAGCATC TTCTAGGCTG AGCTTGACCT TCTCTGCATA GAGAGAGGCT | 14340 |
| ACATGGTTAT CTGTCACAAT GACTACCTTT TGCGGTTGCC AGAGTTCTCG CAACCACTGA | 14400 |
| CCAGCCTGGG CCATACAACC TTTTTCATC TGAATATCAT AAGGATGGTG AGGAATATCG | 14460 |
| ATTCTGATTT TCATAGGAGA GTCTCCCTTT CTTTATTGGT ATTTTCTGT TAAAGACTGC | 14520 |
| CAAATCTCTT CTGTCGGCAT TTCCTTGCCT GTCCACAGTT GAAAAGCTTC TGCAGCTTGA | 14580 |
| TAGAGTAACA TTCCAGACC ATTGACTGCT GGATTGCCCT GACTTCTAGC CCATTTCAAA | 14640 |
| AACGGTGTTC CAAAGGGTTG GTATATGATA TCTGCAACTA AAAGAGTTTC TGGTAAGACT | 14700 |
| ATGTTTTCAG GAACAGGAGA GGATTGGCCA TCCATGCCCA CACTGGTGGC ATTAACCTAGC | 14760 |
| AAATCCGACT CGGCAATCCT TGCTTGCAGT TCAGAAACAT ATTCTAAAGC ACACAAATCC | 14820 |
| ACTTTAAAC CTGTCTGCTC CTGTAACCTG TCTAGGTAAG GTCTTGTTTT TTCCATAGAA | 14880 |
| ACGGAACGAA CAAAGACCGA AATCTGACTG ACGCCATCCA AAATAGCCTG TGCCAAGATT | 14940 |

691

| | |
|---|-------|
| GATTTAGCCG CACCACCTGC ACCCAGCAGG GTCATCTTTT TACCTGAAAT TGTAAGAA | 15000 |
| GGCAAGCACT TAAAAAATCC CTTGCCATCT GTATTATATC CAATTAAATT GCCATTCTCA | 15060 |
| TTGACAACCG TATTAACCGC ACCAATCAAG CGCGCTTCAT CGCTCAGCTT ATCCAAATAA | 15120 |
| GGAATCACCT GCTCCTTATA GGGCATGGAC AGATTGATGC CAAACATCTG GTAGCGACGA | 15180 |
| ATATTGGCCA CTGTTTCTAC CAAGTCACTC GCTTCAATCT CCCAAGCCAC ATAAGCACCG | 15240 |
| TTGGTAGCTG TCGCCTCAA GGCTCTATTG TGGATGAAGG GAGAAATAGA ATGCTTAATA | 15300 |
| GGATTGGCAA CAACTGCAGC TAAACGTGTA TAGCCATCAA GCTTCATCCA AAATCTCCCT | 15360 |
| GATTTTTTTC ATGCTAGCTA GAGAAATCTG CCCAGGGGCA CTAACCTCAT CCAGACTGGC | 15420 |
| AAAAGACCAA CTCGAACCAG TCACATCCGC AGTGATACGA GAGACCTTGC CCACCTTACC | 15480 |
| CATAGAAATG GTCACATATT CCTGTTCAGG ATTGAGGGTT TTAAGCCTC GTGTATAGTT | 15540 |
| CATCAAGTCT AAGACATCCT GCTCCGTGTG AGCCATCACC GCAACCTTAA CAAGTTTGG | 15600 |
| ATTTAGGATC GTCAACTCTG ACAAGATTTC CATCATGTTT TCAGGTGTTT CTTGGAAATT | 15660 |
| ATGGTAATC AAAACAAGAT TTGGGAAGTC CAGCATTTCC TCAAAAACAT CCTTGTAGCT | 15720 |
| ATAGTACTCA AAATCAATAT AGTCTGGTTG ATAGAGTTGC GCAACTTCCT TGATTAGATG | 15780 |
| GATATACTCT TCTGGAGAAA GGTGATTTT TCCACCTTCG GAGCGAGTTC GTAGCGTGA | 15840 |
| AACCAACTCA CGGCCTGCGA ATTTTTCAAA AATGGCTGGA GCTACCTGCA AAATCGCTTC | 15900 |
| TTTAGGCAGA TAGTCGGCAC GCCATTCAAT GATGTCGGCA TCCAGGTACC TCGTGGCATC | 15960 |
| CAGAGCCTGA GCCTCCTCTA AACTTCTTGG CATTACTGAA ACGATTAATT TCATTTACTA | 16020 |
| ACCTTCATAC TAATCACCTT GAGGTAATTA CTACTTTCAT CTTTTTTATT ATAGGCAAAA | 16080 |
| TCTGCTGGAA GACCATATTT GTTAAAAATC TGGTAACTTC TTCCTGCAAA ACCTTTATCA | 16140 |
| ATTTGTCTG TAAATTTCTG ACGGGAACA TTGGCAGCAT TGGTACTGGC AATGATAATC | 16200 |
| CCTCCCGGAT TTAAATCTC AAGACTCTGG GAAATCAACT TGTGATAATC CTTGGCCACA | 16260 |
| GAGAAAGTTT GTTTTTTATT CCGAGCAAAG CTAGGCGGAT CTAGGACAAT CACATCGTAG | 16320 |
| GTCAAGTCTT TGCCTTTGGC ATATTTGAAA TACTCAAAGA CATCCATGAC TATAAACGA | 16380 |
| TGCTCGTCTG TGCTGAGCCC ATTTGCCTGA AAATGCGCTT GAGACAATTC TCGTGAACCT | 16440 |
| TTGGCTAGAT CAACAGAAGT TGTATGGCTA GCTCCTCCCA TGGCCGCAGC TACTGAAAAA | 16500 |
| GCCGCTGTGT AGGAAAACAT ATTGAGTAAG GATTTACCCA TAGCCAAGCC GTCAACTAAA | 16560 |
| CTACCGCGAA CCTCATGCTG GTCTAGGAAA ATTCCTGTCA TCAAGCCATC ATTCATAAAG | 16620 |
| ACTTGATACA GGACACCATT TTCTAAAACA TTGAAAAGT CAGGTGCTTC TTGACCATAA | 16680 |

692

| | |
|---|-------|
| ACATGGGCAG ATTTCATAGTC CAAACCCCTTA AAGCGGATTT TCTCATAAGC TCCTAAAACC | 16740 |
| TCAGGGAAAA CCTGTCTAAA GGCTTCTGAT ATAGTCTGAC GAATCTGATA AACATAAGAG | 16800 |
| TTATACCAAG AAAAGACGGC GTAGTCGCCA TAAAGGTCCA CTGTCAGACC CCCAAAGCCA | 16860 |
| TCTCCCTCTT GATTAAAGAG ACGAAAGGCA GTTGTCAAAT CATCTTGATA GTAGGCGTTT | 16920 |
| CTCTTTTCTT TGGCTTTTCT AAACAACGTT TCAAAGAAAG CTTGATTGAA GGCCACCTTG | 16980 |
| TCTTTGCTGA TAAACCAGCC CAAGCCCTTG TTTTGCTGAG AAAGGTAGGC AGTCCCAAGA | 17040 |
| AAGTTTCCTT CTGACCCTG CACCTCTACT TCCTGATCCT TAAGATTGAC ATTCTCAAGA | 17100 |
| TCCTGGCTT CTAGTAAAAC TAGCCCCCTTA GCAAGCTTCT TTTCAACCCT TTTGCTGACT | 17160 |
| CTTATTCTAT TCATAACTAC CATTATATCA AACTTTTGA CAATTCTCAA AAAAGAACT | 17220 |
| ACCCTTGCTT TTTTACTCTT CTTTAAAAA ATGGTATACT AGACTTCCTG CAAAACCTAGG | 17280 |
| AAGTAAATGT GTAAGAATCA CAGTAAAAA TGCTCTCCG TCTTGAGGA GCATTTCCTT | 17340 |
| TTATCAACGA AAATCAAATA GCAAATATG AAAGTAGCCT CAGGTAACT GTGAGATTAT | 17400 |
| AGGTAGAGAG GTTGTATCAG CAATATGTGT CTGTCAAAT TAGTGACAAA GGTAGTAGAA | 17460 |
| GAAAGATAAA GAAATAAATC AGCTTCAGTA GGTATCTGGA AAATTTGATT TTATAGAGAA | 17520 |
| GCCTTTTGTT ACAAACTCAA TATACTATCA ATAAATAATA TTATAGAAGC AACATAATT | 17580 |
| ATAATTTTAC CTATCTGCAT CATTCTATTT CGAACTCTAA ATATATGTTT TATCAAAAAT | 17640 |
| ACTTGAACA CACACATTAT AGGAATTAAC GTTTTGTAAA TTGAAAAATA TCCAAATAAA | 17700 |
| TAACTATAA ACAACAAAA TAGAACTATG TTATATTTCT TATTCAAAAC ATTCCTCCCT | 17760 |
| ATATATTTTT GATTACCAAT CTTAATCATT TACAACATA TTCTAACAAA CTATAAAGC | 17820 |
| GTTTGTGCGA TTGAATTTAT CAAGCAAGCG ACCAACCAGT TCATCTTTTT TCTATTTCTG | 17880 |
| CCAATATGCG TGACAGGTAA TAATGATAGC CAAAAATAGC AAGAGCAAGC AAGACGATAA | 17940 |
| GAGCTCTTAC TCCCAAGCTG ATGGCAAGGA TAGGGGAGAG AGACTGAACC AAGAATATGC | 18000 |
| TCCCAATTAC AAGGGCCATC AGGATTGCAC TATAAATAAA CAATAAACT ATGGCGACTA | 18060 |
| TGCCATTTGA ACGATTCAAC AGGTCCGTAA TGCTACTCCA ATTGATTGAC AGATTTTAA | 18120 |
| CGTCCTTAAA GTAATGGTGG CAAGAAAGGA TGACACTGGC AATGATCCAG ACTACAAGAA | 18180 |
| GGTAAATCAT CGAAATGATG GGCAAGCCTA GATATAGAGA AAGACCAAGC AAAGTCAGAA | 18240 |
| CTGGTAAAAA GGAAGTGGCA GCATATATAA TCCAAAATTT CACTTTCACA TAACGAGCAA | 18300 |
| AGTCAAAGGG TAAACTCTTA AGAAAATCAA CATTTTCCCT CTCCAAGGAC AAGGCAATTG | 18360 |
| AATGCAGGCT GGTGATATTG TTATTGACAA CTGCTATAAA GAGAGCTATA AAAACAAGG | 18420 |
| GTAACCAGTA TGGAGGATGA ATGTCTGGAA CTATCTGAGA ATCTCGGATT TTGGAATCA | 18480 |

693

GACCGATCAT CATGAGATAA GGAAGGAAAG CACTTGTAAG AAGCACTGTA ATCACGCCAG 18540
 TCCCTGTGCC CAAGAGGGTG AGGTGGTAGC GTAAAACCAT GCGAAAAAAT CCCTTTTGTAG 18600
 TGGTTGAAAT TCTCTCCTTG CTGCGACGTT CTTTTTTGAC CTTCTCCTCA CTATTAAGCA 18660
 GGATCACGTC ATAAAAACGA GGAAGGACCT TCTTTTGGT CAGATAAAGC AGGAAGAGAG 18720
 TTAGTCCTAT CCAAGCGAGC AGACCCACTA AGGCTTCTGT CGAAAAAGGC TCCACTGCTA 18780
 TTTTGTAAGG GATATGAAGA GGATAAAGGA GAAATGGAAT GTCTCTAACT TTGTCAACAA 18840
 TACTTCCAAA AGTCGACTGA AGAAAGAAGA TAAATATTAA AGGTATGAGA ACTCCTATCC 18900
 CAATCATCAC ATTCGAAAAA ATAGACTGAT ACTTCTGAA GACCCTAGTT TGAGCCAAGA 18960
 AATGCACTGC CACTACCATC ACTAGAGCCA CAGAGACAAA TAATAAGGTC AAGGACAGTA 19020
 GCATCAAAGG CAAACCCAGC CATAGAGAAG GAGCTAGCCT AATGTAGAGG ACCAGAAAAT 19080
 AAGCTAGGAT TGGTACAATT CCAGTTAGAG CTGGCAAAAG GACAGACAGT CCTTTAGCAA 19140
 TTATAATCTC TGATTCTTTA AAGGCATAGG GCCTATACGA TACCAAATCC TTACTCTCAT 19200
 AAAAGACATT GTAAAAGGCC GTTAAAGAAG TTGAAAAGGC AATCACTAGT AAAATAGCAA 19260
 TCATCGAGCT AAAATAAATA GGTATTTCTT CAAAAGGAAA ATGAATGGCT ATATTACTAA 19320
 AACAGATGAT CATCAAGAGA CTGGAAAAAA TGTAAGAACT TAAGACTCTA GCGGAAACAT 19380
 TTACTTTTTT 19390

(2) INFORMATION FOR SEQ ID NO: 87:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 18436 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 87:

CCGAGCGTCG TTACAGACTT TATCAAGATT GGACGCAAGA AGAAATTCAA CATATAAAGG 60
 AAAATATGGC ACAATCTCCA TGGCATACTC ATTACCATGT TGAGCCAAAA ACAGGACTTC 120
 TCAACGACCC AAATGGCTTT TCTTACTTTG ATGGCAAGTG GATCCTCTTT TACCAGAATT 180
 TTCCTTTTGG TGCAGCCAC GGTTTAAAT CTGGGCACA GCTAGAAAGT GATGATTTGA 240
 TTCCTTTTAA AGAACTGGA ATCAAAGTTT TACCAGATAC TCCATTAGAT AGCCACGGTG 300
 CCTACTCTGG TTCTGCCATG CAATTTGGCG ATAATTATT CCTATTTTAT ACAGGAAATG 360
 TTCGCGATAA AAAGTGGATC CGTCACCCAT ACCAGATCGG TGCTTTGATG GACAAGGAGG 420

| | | | | | | |
|------------|------------|------------|------------|------------|-------------|------|
| 694 | | | | | | |
| GTAAGATTAC | AAAGATTGAC | AAGATCTTGA | TTGACCAGCC | AGCAGACTCT | ACTGACCACT | 480 |
| TCCGCGATCC | ACAAATTTT | AACTTTCAGG | GTCAATATTA | TGCCATTGTC | GGCGGACAAG | 540 |
| ACTTGGAGAA | AAAAGGTTTC | GTTCGTCTCT | ACAAGGCTGT | CAATAACGAC | TACACAAACT | 600 |
| GGCAAGCAGT | TGGCGACCTT | GACTTTGCTA | ACGACCGTAC | TGCCTACATG | ATGGAATGTC | 660 |
| CTAATTTGGT | CTTTGTAGAG | GAACAACCTG | TCCTTCTCTA | CTGTCCACAA | GGATTGGATA | 720 |
| AGAAAGTTCT | AGACTACGAT | AATATCTTTC | CAAATATGTA | TAAGATCGGG | GCTTCCTTTG | 780 |
| ACCCATAAAA | TGCCAAAATG | GTAGATGTGT | CTCAACTTCA | AAACATGGAT | TACGGTTTCG | 840 |
| AAGCCTATGC | AACTCAAGCC | TTCAACGCTC | CTGATGGGCG | TGCTCTAGCA | GTTAGCTGGC | 900 |
| TTGGTTTGCC | AGATGTTTCT | TACCCATCTG | ACCGTTTTGA | CCACCAAGGA | ACCTTCTCTT | 960 |
| TGGTCAAGGA | ACTCACTATC | AAAGACGACA | AGCTCTACCA | GTATCCAGTC | GCTGCTATTA | 1020 |
| AGGACCTTCG | TGCTTCTGAA | GAAGCCTTCT | CAAACCGTTC | CCAAACCAAG | AACACTTACG | 1080 |
| AACTTGAAC | CAACTTGGA | GCTAATAGCC | AGAGCGAGAT | TGTCTTACTT | GCTGATAAAG | 1140 |
| AAGGTAAGGG | ACTTTCAATC | AACTTTGACC | TTGTAAACGG | TCAAGTAACA | GTGGATCGTA | 1200 |
| GCCAGGCTGG | AGAACAGTAT | GCCCAAGAAT | TTGGGACAAC | TCGTTCTTGC | CCTATCGAGA | 1260 |
| ATCAGGCTAC | TACTGCTACA | ATCTTCATCG | ATAACTCTGT | CTTTGAAATT | TTCATCAATA | 1320 |
| AAGGAGAAAA | AGTATTTTCT | GGTCGTGTCT | TCCCACATGC | GGACCAAAAT | GGTATCTGTA | 1380 |
| TTAAATCTGG | AAACCAACT | GGAACCTACT | ATGAATTAGA | TTATGGTCGC | AAAACCTAACT | 1440 |
| GATGTCGCCA | AACTTGCAGG | CGTCAGTCCT | ACTACCGTTT | CTCGGGTTAT | CAATAAAAAA | 1500 |
| GGGTATCTAT | CTGAGAAAAC | CATCCAAAAA | GTCAATGAAG | CCATGCGAGA | ATTGGGCTAT | 1560 |
| AAACCAACA | ACCTGGCTCG | TAGTCTGCAA | GGAAAATCAG | CTAAGTTAAT | CGGCTTGATT | 1620 |
| TTCCCAATA | TTTCCAATGT | TTTCTATGCA | GAATTGATTG | ATAAATTGGA | ACACCAACTC | 1680 |
| TTCAAAAATG | GTTACAAGAC | CATCATCTGC | AACAGTGAAC | ATGATTCTGA | GAAGGAACGC | 1740 |
| GAATACATCG | AAATGTTGGA | AGCCAATCAG | GTGGACGGCA | TCATTTCTGG | TAGTCACAAC | 1800 |
| CTAGGAATCG | AAGACTACAA | TCGTGTGACA | GCGCCGATTA | TTTCCTTTGA | CCGAAACCTA | 1860 |
| TCGCCAGACA | TCCCTGTCTG | CTCCTCTGAC | AACTATGCTG | TGCGGGTTCT | TGCTGCCCAA | 1920 |
| ACCTTGGTCA | AGACAGGTGC | CCAGTCTATC | ATCATGATTA | CAGGGAATGA | CAATTCTAAT | 1980 |
| TCGCCAACCG | GACTGCGCCA | CGCTGGTTTT | GCATCCGTAC | TCCCAAAAGC | TCCTATTATC | 2040 |
| AATGTTTCCA | GTGACTTTTC | TCCCGTCAGA | AAAGAAATGG | AAATCAAGAA | TATCTTGACC | 2100 |
| CGGAAAAAAC | CAGATGCCAT | TTTTGCTTCG | GATGATTTGA | CAGCTATTCT | GGTCATTAAA | 2160 |
| ATCGCTCAAG | AATTGGGCAT | TTCTGTCCCA | AAAGAGCTCA | AGGTCATCGG | CTATGATGGG | 2220 |

695

| | |
|--|------|
| ACCTACTTTA TCGAAAATTA CTACCCCTCAA TTGGCTACTA TCAAGCAACC TTTGGAAGAG | 2280 |
| ATTGCTTGTC TCACTATTGA TCTTCTCTTG CAAAAGATTG AAGGCAAGGA AGTCGCCACA | 2340 |
| ACTGGTTACT TCTTACCAGT TACGCTATTA CCAGGAAAAA GTATTTAAAC ACAAGAAAAAC | 2400 |
| TCAGACCGAT TCGTCTGAGT TTTTATGATC TTAAATTTTC GAGATAGCGC TGGGCTGTCT | 2460 |
| CTAGGTAA GGTTTTATCT GAGATGAGGC GCTCTACTAG GGGAGCAACT TCAGATTAC | 2520 |
| TAGCCCCAGC TAGGAGAGCT AGGGATTTGG CCTGTAGTTT CATGTGGCCT TGCTGGATGC | 2580 |
| CCGTACTTAC CAAGGCTTTG AGGGCTGCAA AATTTTGAGC AAGACCGATG GACACGATAA | 2640 |
| TCTGGGCTAA TTCTCTGGCA GAAGGATTTT CTAGTAGATC ATGACTGAGA ACTACACGTG | 2700 |
| GGTTGAGGCC GATAGAGCCA CCCTTAGTCG CTACAGGCAT GGGCAGGGTC ATCTCACC GA | 2760 |
| CCAATTTCTC TCTTTCAAGG TCCAGCGTCC AGCAGCTAAG ACCTTGATAG CGTCCATCTC | 2820 |
| GACTGGCAAA GGCATGGGCC CCAGCTTCGA TGGCACGCCA GTCATTACCA GTGGCAATCA | 2880 |
| AAATCGCATC AATACCATTA AAAATTCCTT TATTATGAGT AGCAGCTCGG TAAGGATCAG | 2940 |
| CCTGCGCAA CTGACTAGCC AACGCAATTT TCTCCGCAAT CTCTCGTCCT TGATCCTTTT | 3000 |
| GGCGGCTCAA GTAGCGAAAG GCGATGCGAC AGCTTGCACT CACCAGAGAA TCGGTGCGCT | 3060 |
| AGTTGGACAG GATTTCCATG AGACTCTGTC CCTGACTGAG TTCTTCTAAG ACTGGTTTCA | 3120 |
| AGGCTTCCAG CATGGTGTG AGCATATTGG CACCCATGGC TTCCTGGGTA TCGACATGAA | 3180 |
| TATAAACAAC GAGAAAGTCT GGTTCGCCTT TTATCTGCTC GACATGCAGA TCACGCGCCC | 3240 |
| CACCTCCACG TTAAACGATA GAAGGATAGG CTTGATTGGC AAGCTCCAAG AGCTCCGCTT | 3300 |
| TCTTGCTGGC AATCTTCTCT TCGCTAGTT TAGGATTAGC AACTTGATAA AGGGCTACCT | 3360 |
| GCCCAATCAT CTGTCGCTGA TGGACTTGTG CAGTAAACC ACCTGCACGC TTGATGATTT | 3420 |
| TGCTGGCATA GCTGGCCGCC GCAACCACAG AGGGTTCTTC TGTCACATAG GGAACGGTGT | 3480 |
| ATTCTGACC GTTGACAAGT ACCTCCGGAA CCAAGTGAATA AGGCAGAGAA AAAGTTCCCA | 3540 |
| CTACATTCTC ACTCAGCTGG TCTGCCACAG TCACGCTCAT CTGTTTATCC TTCTCCAGAC | 3600 |
| TAGCTTGTCT CTCAGGACTA AGGAGCGCCT GAGCTTTTAA CAGCTCGAGG CGCTCTTGGT | 3660 |
| ATGATTTTTT AGAAAATCCA TTCCAACCTA TCTTCATTAT TTTTCAACCT TGCTATAACG | 3720 |
| GCCTTGGTGG TCGAGAATTT CAACCAAGGC AAAATCTTGA TTTTCATAGC CAGCAAACTG | 3780 |
| GGCAGAGTTA GTTTCATCCA AGTTTACTTC CTCAAAAAAG ACCTTTTCAT AGTCTGCAAC | 3840 |
| GGATAGGGCA GTTCGTTGGT TGAGCTTGT CAAACGGTCT TTATCCAAAT AAGCTTCATA | 3900 |
| TCCTTCAACC AATTCACCAC TGAAGAACTC AGCCACAGCT CCACTTCCGT AACTATAAAG | 3960 |

696

| | |
|---|------|
| GGCGATTTTA TCCCCAGCTT TCAAGCTATC TGTATTTTCC AAGAGAGACA AAAGTCCAAG | 4020 |
| GAAAAGTGAA CCTGTGTAGA TATTCCTCAC CTTTGTACTG TAGAGAATAG ACTGGTCAAA | 4080 |
| ATGCTTTTGT AAGAGGTCTT TTTTCTCTTG AGGCAGGCTC TTATCCATGA TTTTMTCAA | 4140 |
| GCCTTTTAGC GCTAATTTAG GATAAGGCAA GTGGAACAA ACAGCCGCAA AATCATCCAA | 4200 |
| AGTAAGCTGG TAGCGTTTTT GATATTCAAG CCAAGTCGTT TTCAAACAT CCAAGTATTG | 4260 |
| TTGGGTAGAA TAGACACCAT TTACATAAGG AGTTGTCGAG TAATTTGGTC GCCAGAAATC | 4320 |
| CATGATGTCA CGGGTCTGAG CTACATTGTC ATTATTAAAG GCCATCATGC GTGGATTTTG | 4380 |
| TGTAATCAAC ATAGCTACAC TTCCAGCACC TTGAGTTGGT TCTCCTGGAG TTTCAATACC | 4440 |
| GTATTTGGCA ATATCACTGG CAATGACCAA GACCTTGGAC TCCGGAGAAT TTTCCACATG | 4500 |
| CAATTTGGCA TAATGGAGGG CAGCAGTCGC TCCGTAGCAG GCTTCTTTAA TCTCGAACT | 4560 |
| ACGAGCAAAG GGCTGGATGC CCAGCAAGCC ATGCACAAAG ACGGCCGAG CCTTACTCTG | 4620 |
| GTCAATTCTT GACTCGGTCG CCACAATGAC CATGTCAACT TCTTGTCTTT CTGCTCAGT | 4680 |
| TAAAATAGAG TCACTAGCAC TGGCCGCCAA GGTACGATA TCCTCAGTTA GGGGCGCAAT | 4740 |
| ACTCAATTCC TTGAGTAAGA GTCCTTTACT TAATTTTCA GGGTCAATTC CCCTCGCTTC | 4800 |
| TGCTAAGTCT TGTAATTTCA AGACATATTG ACTGGTCGCA AAACCAATCT TATCAATACC | 4860 |
| GATTGTCATA TTTACCTCTG TTTTATCATT CATGTAAAA ATCGTTCTAT ACTATTTTAT | 4920 |
| CACAAATGGC AGTAAAGAG AGAAAAAGA CTTGATTAC CAAATCAAGC CTCTTATTGG | 4980 |
| TCATCATTTT AAAGAATGAT TAGTTGCTAG AGAGTTCACC GATATAAGTA GCTTTATAAG | 5040 |
| CTCCATTAC AGTTATCAGC TCCTGGAGGA TCAAATTTCC TGAGTAAGTC CTCCCACCT | 5100 |
| CATCTACAAA TTTTGTATAA AACTGACTGG TCGGAATTC TCTGACATCC TTATCAAATG | 5160 |
| TCTTATCAAG TGTTTTACTA ACCTTCTCAG CAATCAATTG ATGCTCTTGC CATCCACTTT | 5220 |
| GAAACTCTGA GCCCGAATA GAAACCATGA CTGGGATAAA CAACAAGGTC AGTAGATTTA | 5280 |
| CAGACAATAA GGAAAGTAGT AGACTTCCTG CAAACTAGA ATCCTAGTTC ATGATTGATA | 5340 |
| ATACCAGCAA TCAAATTCAT TCGTAATCCG AAGCGTTTAC GATGATTTTC ATAGGTTGTT | 5400 |
| GAAAACATTT TAAACGTTTT TACTTTGGCA AAGATGTTCT CAACCTTGCT TCTCTCCTTA | 5460 |
| GATAGCGCAT GGTACAGGC TTTATCTTCA GCTGTTAGCG GCTTGAGTTT GCTGGATTTA | 5520 |
| CGTGGAGTTT GTGCTTGAGG ATATATCTTC ATGAGCCCTT GATAATCACT GTCAGCCAAG | 5580 |
| ATTTTACCAG CTGTCCGAT ATTTCTGCAA CTCATTTTGA ACAACTTCAT ATCATGACTA | 5640 |
| TAGTTCACAG CGATATCCAA AGAAACAATT CTCCTTGAC TTGTGACAAAT CGCTTGAGCC | 5700 |
| TTCATAGCGT GAAATTTCTT TTTACCAGAA TCATTCGCTA ATTGTTTTTT AGGGCGATTG | 5760 |

697

| | |
|---|------|
| ATTTTACTT CCGTCACATC AATCATATC GTGTCCTCAA AGCTGAGAGG AGTTCCTGAA | 5820 |
| ATCGTAACAC CACTTTGAAC AAGAGTTACT TCAACCCATT GGCTCCGACG GATTAAGTTG | 5880 |
| CTTTCGTGGA TACCAAAATC AGCCGCAATT TCTTCATAAG TGCGGTATTC TCGCACATAT | 5940 |
| AGAAAGCGTT ATCAATTTAT TTATCTCATT TTTCAGAAAA TTCTTTTATT TCTGTAAAGT | 6000 |
| CTACGATACT CGATGTGTTT TTATATAATG ATAGAGTCTG AGAATCACTG TTCCGCTAGC | 6060 |
| CATTCCAATA GAGATTACCA AAGCCAACAT GACAACCAAG GTCGCACTTG CCAGTGCTTT | 6120 |
| ATTATAGTCC CCTGTCACAA AAAAGGCAGT TGTTCGGTAG GAGAGATAAC CTGGAACGAG | 6180 |
| CGGTGCCAAA ATGGCCAAGA TAAAGACCAC AGCAGGTGTC TTATAAAGAA TACTTAAAAAT | 6240 |
| CTGGCTGACA CAAGAACCAA TAATGGCTGC AATGAAGTA GCTACAATGA CATTTGGTCGG | 6300 |
| TTCCTTGAGC AAGAGATAGA TTAGCCAGAC AGTCATGCCC AAAATCCCTC CAGGTAAGAG | 6360 |
| CATAGACCGT TGCACATTGA GTACGATTAA AAAAGTGATA ATGGCAAGAA AACTTGCTAC | 6420 |
| TGCTTGTAAT AAAAAGGTTG TTAGTGTCAT ATTAGTTCAT CAATACCAAG GCGACAGAAG | 6480 |
| TTCCTGCCCC TAAAGCGAGG GTAATGAGCA GGGATTCAAA CATCTTACTC ATACCAGAGT | 6540 |
| TTATGTGGTT GGTACATAATA TCACGGACCG CATTTGGTCAA GGCAATACCT GTTACAAACG | 6600 |
| GCATGACCGC ACCAGCTATA ATCAAATCTG CCGTTGAAGG AAAACCTGTG TAGCGAGCCC | 6660 |
| AAACTGGGC AATTATCCCA AAGACAAAGG CTCCAGCAAA GGCTGTCACA AAGGGAATTC | 6720 |
| GGATAAATTT TTCCACATAG AGGGAAAAGG CAAAACCAAA TAAGGTCGCC ACTCCTGCCC | 6780 |
| CAAGTGCGTC GTAGATATTT CCGCTAAACA TAACTGAAAA GAAAGGAGCA CTAAAGGTCG | 6840 |
| CAGCCAGAGT TACCTGCAAC TTAGTATAGG GAAGGGGTTT GGCTTGCAAG GCCGTCAATT | 6900 |
| GCTTAAAGGC TGTTCCTAAG TCAATCTGCC CCCCAACTAG CTGACGAGAA ATCTGGTTCA | 6960 |
| CATCGCAGAC TTTTTCGATG TTATAAGAAG AGGAGGTCAC GCGCTTCATG CGCAAATATT | 7020 |
| GGTATTTTCA ATAGAGAAAA AGATAGCGGC AGGCATGGCA AGGACATTGC AATCCACAAT | 7080 |
| CCCCTGCGAA TGCGCGATTG GAATCATGGT ATCTTCTACA CGATGGATTT CTGAGCCACT | 7140 |
| TTTAAGGAGA ATAGTCCCCG CTAGCATAAT CACATCAATG ACGGCATTTA ATTCTTTTGA | 7200 |
| TTCTTCCATG CTTTCCTCCT TTTATCAACT CCCTCTATTC TATCACAAAT CCGGACTCAA | 7260 |
| AAAAAATCTT TGCCATGAAA TCATGACAAA GATTGATTAC TCATTTTGAT TATCCATCTG | 7320 |
| CTTTTAAGGA GTAGCTGAAG TTGTTTAGG TTTGTAGATT GAAATCTTGA CTCTAGTCTT | 7380 |
| ATTGAGGTCT ACCTTTTCAC CTGCTCTAGG ACTTTGTTCA ACAACCATGC CTTCTGCACT | 7440 |
| ACCTGCAGGC GCTGTCGTCA CTTCTACAAC TTCTATATTA GCTTCCTTAA TCCCAACAAT | 7500 |

698

| | |
|--|------|
| TTGAATCAAA TTGTTCTTAG TAAACTCCAA GCTAGAACCA ATGTAACCTCG GCATGGCAAC | 7560 |
| ACTTGTAAGT TTTTGTAGCTA CTGTCAAGAC AATTGAGTA GGTTTACTCA CATCATAAGT | 7620 |
| CGTTCGGCA CCTGGACTTT GTTTCATAAT CGTTCCTGGT TCGCTTTCGC TGGACTCTTC | 7680 |
| TTCTCTATC TTAATCAAAT TCTCAGGAAC CTTCTTCTGC TTGAGTCTG AGATTACTTC | 7740 |
| TGTAGAGTTC CGTCCAATAT AGTTCCCTAA TTGAATCGTC GTAGCTTTTT TAGCTACTGT | 7800 |
| CAAAACAATT TGAGTTGCCT TGCTCAAGTC ATAGGTCGTA CCTTCTGGTA GACTTTGCCT | 7860 |
| CAGGACCGTT CCAGCCTCAC TCTCATTCGA CTCTTCTTCC TCAATTTTAA TCAAATTATC | 7920 |
| TGGAACTTTT TTCTCTTTTA ATTCCGCAAT GACATCAGAG GATTTCCGAC CGACATAATT | 7980 |
| ACTAATTGGA AAAGATTGCT TGCCTGATGA GACAACCAAA TTGATTTTCG TTCCTTCTTT | 8040 |
| TCGACCAGTT CCAGCGCCAG GATCTGTACG GATAATCCGC CCTTCTTCCA CCTTTTCACT | 8100 |
| AGCCTCTGTC TTCTCCTCAC CAATCTCAAA ATTGGCTTTT TTGAGCGTTG CCTTGGCCTC | 8160 |
| TGCAACTGTC TGACCTGCCA CATCTGGAAT GGCAATGGTT GCAGGAGTTC TGGATAGTAT | 8220 |
| CCAAATAAGA GAAGCTGCCA CCAATACAAG GCTGGCCAAC AAAATCAGGT AACGCATCTT | 8280 |
| AAATCTATGT TTTTTCGGTG CTTGTGGTTG GTAAGTTTCC TCTGTCACAG CCTGGCTTGG | 8340 |
| GTTTTTGATT GATTGTGTT CTGTTTGC GC TTGAACCTTA GGAATAGATG TCAAGGTAAT | 8400 |
| CTGAGAAACC TTCGGCAAGG TCTTGGTATC TGCCTTGCTC GTTTCATCAA AGATTAACTT | 8460 |
| ACTTTCATTT CTACGATTGT AGGACAAGCT ACTAGACAAG TCCACATACA TCTCTGAAAC | 8520 |
| CGAGCGGTAG CGATTGGTCA ACTTTTTAGC AGTTGCCTTG ATAATAACAT TTTCTAAAGC | 8580 |
| CTGAGGTACA GATGGATTTT CTGCAATAAC GGACGGCAGG GGTTCCTGGA AATGCTGGAG | 8640 |
| GGCAATGGTC ACCGCGCTAT CCCCCTCATA AGGGATATGG CCTGTCAGCA TCTCATAGAA | 8700 |
| AATAATCCCC ATGGCATAGA TATCACTCTG CACAGTCGCC TTCGAACCAC GCGCCTGCTC | 8760 |
| TGGTGACAAG TAATGAACTG AGCCCAACAT CGAGTTAGTC TGGGTCAGAC TTGTCTCTGC | 8820 |
| AAAGGCTACA GCAATCCCAA AGTCTGTGAC CTTGGCAGTC CCATCTGGTG TCAAGAGGAT | 8880 |
| ATTTTGAGGT TTCAAGTCCC TGTGAACAAT TCCTCGAGTA TGGGCCAAGC GCATAGCCAA | 8940 |
| GAGAATTTGT CCCATGATAC GGACTGCTTC TTCATTAGAA AGAGGATAAT GTTCCTTGAT | 9000 |
| ATAGCGTTTG AGGTCCAGTC CAGCCACATA CTCCATAGCT AGGTACTGTT GACCGTCTTC | 9060 |
| CTCGCCAATA TCTGTTATCC GAACGATATG AGGATGGTCT AGATCTGCCA TAGCTCTCGC | 9120 |
| TTCACGCTGA AAACGAGCTA CAGCTATCGG GTCCGCTCTG TAGTTGGTCC TCAGAACCTT | 9180 |
| CACTGCCACT TCTTCCCAT CTAAGATTAA GTCTTTGGCT AGGTAGACAT CCGCCATACC | 9240 |
| TCCTCGACCA ATCTGTTTGA CAATCCGATA GCGTCCGGCA AAAATCTTGC CGATTGGAT | 9300 |

| | |
|---|-------|
| CATTCTGCAT CCTCCTCGTT CATAGAAACA AGGGCAACCG TAATGTTGTC TAAACCTCCT | 9360 |
| GCATTGTTAG CAAAACGAAC AAGTGTCTCC GTTTTATCTG CTAAAGGAAT ATCACTGGTT | 9420 |
| ACAAATATCAC GAATCTCACT GCCTGAAATC ATGTTGGTCA AGCCGTCCT ATTGAGCAAG | 9480 |
| AGATAGTCAC CTGACTCAAG GATAACTGTC CCAAAATCAG GCTGAATTC ATCTTTTGC | 9540 |
| CCAATAGACT GGGTGATAAT ATTTTGTG GGATGAGCTT CTGCCTCTTC TGGTGTCAAT | 9600 |
| TGACCAGCCT TGAGCAATTC ATTAACCAAG GAATGATCGC TCGTCAACTG ATGGTATTCT | 9660 |
| TCTCCACGAA TCAAGCCGAT ACGCGAATCA CCAATATGAG CATAGATAGC CTGATTATCA | 9720 |
| ATAATAGCAA GGACTTCCAA AGTAGTCCC ATGCCTCTGT AAGCTTCATC CTGACCAAGC | 9780 |
| TGGTGAATCT TTTGATTTC AATTCTAGG TAATGGGCGA ACCATTCACG CACTTCATTG | 9840 |
| ACTGTATCGA TCTGGGTATC AACCAGCT ACACCCAGGT CTGTGACCGC CATTTCACTA | 9900 |
| GCGATATTCC CTGCGCGATG ACCTCCCATC CCATCAGTA AAATAATCAT GGTACGTCCA | 9960 |
| GCTCTATTGA CATAGTGGTT GACATAGTCT TGGTTATTG TTCGTTCTG ACCAACATCT | 10020 |
| GTTAATAATG AAATTCCAT GTGTCAGTTC CTTCCCTAATC CGATATCTTG CGAAATTGAC | 10080 |
| TGATGAAGAA TCCATCACTT CCATACAATT CAGGTGTAAT GAGGATACAG CCGTCTTTCA | 10140 |
| TGATATCCTT ACATTCATGT TCTAGTTTA CCTGCTCGAA CTCGGGATGA CTCTCTAAAA | 10200 |
| AGGCCTTAAC GACTTGAAAA TTCTCCTCTG AGACGATAGT GCAGGTGCTA TAAGTTATTA | 10260 |
| TACCACCTTT GCCTAGTATT TGACAAACAC TACCTAATAT TTCTAACTGA ATTCCTGCA | 10320 |
| AGGACGCGAA ATCTGCCGTT TCTTTATTGT ATTTGATATC TGGTTTTCGG CGCAAAAGAC | 10380 |
| CGATTCCTGA ACAAGGAGCA TCCACCAAAA TCTTATCAAA GGAATCCTGG TCAAAAAACT | 10440 |
| CATGCACCTT TCTGGCATCC AATTTTGTAG TTTGAACCG ATCTTCAACT CCCAGACGTT | 10500 |
| GGGCATTTTC TTGAATTAAA TCCAACCTGT GGTGCTACAA GTCCAGAGCA GTAACCTGAC | 10560 |
| CTGTCGTAAG ATAAGAGGCT ATATGGGCTG TTTTCCACC TGGAGCCGCA CAGGCATCAA | 10620 |
| GCACTCGCTC ATCACCTTGT AAATCAAGCG TCGGAGCAAC CAGCTGACTG GACTCATCTT | 10680 |
| GGATGGTAAT GGCTCCATCC GCAAACAAAT TATGCCCTGC AAAATGCCCC TGCTCCTTAA | 10740 |
| CCAGACCAGT GGTGCTAAA AGGGAATTAT TCGCCTCAA CAAGGCTTGG ATTCCTCTT | 10800 |
| TTGACTTAG GTCTGTTACA CGAATACTGG CTTTGTTCG CACTAACAGG CTTTCAAAGA | 10860 |
| TGGCTTTTGC TCTCTCCTCT CCGTATTCTT CCTTGAGTTT GGCAACTAGC CAACTGGGA | 10920 |
| GAGAATAGGC AATGGAGTCA CGCTTGTTTT TTCGCTTGAT GCTAGCAATA TCTGGCCAGC | 10980 |
| CTTCACGCAA GATACGGCGA AGGACAGCGT TGACCAATTT TTCACTGCCT TTTTACGGA | 11040 |

700

| | |
|--|-------|
| GTTTGGCCAA TTCCACTGCT TCATTAACCA CAGCATGATC TGAATCTTG TCCAAATAGC | 11100 |
| GGAGTTGGTA GGCAC T CATG AGAAGAAGGA CATAGAGCCA GCTGTCTAAC TGGTCTCTGT | 11160 |
| CTTCGATAAA GTGGGATAGG TACCATTCCA GAGTCAGTTT ACGGGCTACC GTTCCATAGA | 11220 |
| CCAGCTCGGT CACTAAGCCC TTGTCTGCTG CCAAAAGTTG ACTTCCCTTT AGATGCTTAT | 11280 |
| TTAAGGCGAT ATTTGAATAT GCTTGGTTCA CAAAAACATC CTCTAGCACT GCTAGAGCTA | 11340 |
| AACTTCTAGC CGTTTCTACT TTAGTCACCA AATCGTTCTC CTACAGTCAA TGTACGTCCA | 11400 |
| ACTCCGTTGA GGAAGGAAGC AATGTCCATC TTAGGCTTAC CAGCTGGCTG CACTTGTTTG | 11460 |
| AGGGATAGAG CCCCTTCAGC CGTTGCGACA ATCAATTCTT TCTTGCCGAT AGAGAGAATC | 11520 |
| TCACCTGGAT TTCCCTGACC TTCTACTGGT AGGCTTCAT AAATCTTAA GCGGTCGCCC | 11580 |
| TTAAGGAAAG TATGGGCAAC AGGCCAGGGG TTCATTCCAC GAATTGGTT AAAGAGTTGA | 11640 |
| CGATTGGTTT TGTTCAGTC CAGTTTCTT TCCTCTGGCT TTATATTTGG AGAGAAGGTA | 11700 |
| ACCTGACTCG TATCCTGCGG TTCAGGTTTG ATATCACCAG CAATATAGGC AGGCAGAGTG | 11760 |
| TCCAAAAGCA AATCAGGACC AACTAGCGCC AATTTTTCAC ACAAGGTGCC AACATTGTCC | 11820 |
| TCATCTGTGA TCGGAATGCT GCGACGAGAA ATCATATCTC CTGCATCCAT TTCCTTAACC | 11880 |
| ATTTCCATGA TGGTCACACC AGCTTCCTCA TCCCCTTGAA TCAAGGCATA ATGGATAGGC | 11940 |
| GCACCACCAC GGTGTCTAGG AAGGAGGGAG GCATGAACGT TGACAGCAAA GTCCATGCTA | 12000 |
| TCAAGGAGTT TGCTTGGGAG AAAGTGCCCA AAAGCAGCAG TCACAATTCC ATCTGCTCCT | 12060 |
| AGCTTCATAA GATCTTCCAT CTCTGGACTT CCAGATAATT TTTCAGGTTG GTAGATAGAT | 12120 |
| AGTCCTGCTT CCTTGGCAGC CTGCTTGACT GGGGT T TCTT GGATAACTTT TTTACGACCA | 12180 |
| ACAGCACGGT CTGGCTGGGT CACAACGGCT AGAATTTCTG AACGGTCATC TGTCAAAAGT | 12240 |
| CCTTTTAAGA CTGTTGCTGA AAAGTCGGGG GTCCCCATAA AGATTAGTTT TGTCATATCT | 12300 |
| TCTCCTTCTT ATAAAAATTG CTGCGGCTCA TGGTCAATGC TGAGACGGAG CTCACTATTT | 12360 |
| TCCCGTTCTT GAGTCAAGGC TAAACCTGG TTGAGGGTCG ACCCCAGCTC ATCTTCTAAA | 12420 |
| CGGTATTTAA TTAAATCTG GTAATGATAG AGTTGTGGG TACGGGCAAT CGGTTTGGC | 12480 |
| GTGGGCCCCA GAATGGGACT GGTCTCTGAC AAGCCTGACC GCAAAATGTT CATGACTTCA | 12540 |
| TAGGCACGTT TGAAAACCTC TTCTTCTTTC TTGTGAGAAA GGGTAATACC AATCGTGAAA | 12600 |
| TAGTAAGGTG GATAGCCGAG TTGTCGCTG ATTCCCATTT CATAGGCATA AAAGCCTTCG | 12660 |
| TAATCTTGAT CCTTGGCAAA TCGAATAGCA TAGTGCTGCG GATTGTAGGA CTGTATCAAG | 12720 |
| ACTTGACCTG CCTTTTCAGC CCGACCTGAT CGACCTGCCA CCTGAGTCAA GAGCTGGAAG | 12780 |
| GTCTCTCAG AAGAACGGAA ATCAGGCAGA TTCAAGGCCG TATCCGCATT TAGAACTCCG | 12840 |

701

| | |
|---|-------|
| ACTAGGGTAA CATTGGGAAA ATCCAAACCC TTTGCAATCA TCTGAGTACC AAGTAAAATA | 12900 |
| TCCGCTTCCC CTCGCCCAA CTGGTCAAGC AAGGCTTGGT GACTGCCTTT CTTTCGAGTC | 12960 |
| GTATCCACAT CCATCCTCAG AATGCGAGCT TGGGGAAAGA GTTCTGCTAG CTCATCATAA | 13020 |
| GCCTTCTGAG TTCCCGTCCC ATAGTAACGA ATACTGCGGC TCTTACAGTT AGGACAGACC | 13080 |
| TGAGGAATAT CCTTCGAGAA ACCACAATAA TGGCAGTTCA TAGTCTTGGT ATCCATATGC | 13140 |
| AAGGTCAGAG AATATCGCA GTTGGGACAA GTATCCACCG TCCCACACTC CCGACACATG | 13200 |
| ACAAAGCTAG AATAACCACG GCGATTGAGC ATGAGAACCA CCTGCTCTTT TTTAACCAGA | 13260 |
| CGGTCTTGGG TAGCCTCTAG CAAAGGAGGC GTAAAGTTTG ACGTCTCATT TTGTCCGATA | 13320 |
| TAGTCTCGAA AGTCAATCAC TTGAACCTCA GGGATTGTAG CCAAAGGATT GGCACGTTGG | 13380 |
| GTTAGACGTA AGTGTGATA GACGCCTTTG CCAGCACGTG CCCGGCTCTC TAAGCTCGGC | 13440 |
| GTTGCAGATC CAAGTACCAG AGTTGCTTGA TTATACTGAG CCCGTAAAAT AGTACCTCT | 13500 |
| CTGGCATGGT AACGGGGATT GCTGTCCTGC TTATAAGCCG CTTCATGCTC TTCATCAATA | 13560 |
| ATCATGACAC CCAGATTTT CAGAGGAGCA AAGATAGCAG ATCTGGCACC AACAACAACT | 13620 |
| TGGGCATCGC CACGCTCCAC CTTGCGCCAT TCATCATACT TTTCACCATT GGATAATCCT | 13680 |
| GAGTGAAGAA TGGCTACCTT GTCCCCAAAA CGTGCTATAA AACGCTCGGT CATCTGAGGA | 13740 |
| GTCAAGGAAA TCTCAGGTAC CAGCAAAATA GCTGTCTTGC CCTTATCCAG GGCACCTTGG | 13800 |
| ATAATCTGCA AGTAAACCTC GGTCTTCCCA CTTCTGTAA TCCCTTGAAG TAGAAAGGGA | 13860 |
| GGTTGAGAAC TGCCAATAGA ACTCACAACC GCATCACCGC CCTGTCTTTG TTCTGGATTT | 13920 |
| AACTCCAAAG GTCTACTTGC TTCAATTCCT TCAAAATAAG CAGCCGAGCG TTGAACTTCC | 13980 |
| TTTTGGACTA TGGTAACAGC ACCTTGATCC ACAAAGAAGT TGACTTGCTC TCGCGAGTAG | 14040 |
| GACTCTAACA AGCTAGCCAA GGAAGCGCTC TCTGGATGAG ACAGCAGATA ATCTCTCAGT | 14100 |
| TCCAACTTTT TCTTGGCACG TGTAGAAATC TCAACACCTT CTAATTGAGC ATGGTCAACC | 14160 |
| TCATACCAAG ACTGGGTCTT GACCTTCTTT TGATCGACTG CCTGATAATC CAGACCAAGC | 14220 |
| AGGCCTTTTC TAGTCAAACG CATCATTTCA GCTTGCTTGG CAAGGTCTAG TGAAGAAAAG | 14280 |
| GCTAGCGAAT CTTCTGAACC AAACAGGCGC ACTCGTTCTT CCTGACTCAA GCCTTCCAGA | 14340 |
| GGATAGAGAA TCTTGTGATA GCTAGAATTC AGAAACCCTG GAAGCATGGC CTTGAGGATA | 14400 |
| GAGATTTTGT AGGAGAAGAC AGATTTGCGT AACTCCTCAG CCAGCCAGAG TTGTTCTGGC | 14460 |
| GTGAGAACAG GAGAAAAATC CAGCACCTCT GCAATATCTT TTAAATCTTG CTCCATCTCT | 14520 |
| TCTCCATCTG ATTGGGACTT CAAACCAAGA ACAATCCCTT GAATCAGGCG ATTACCCTTA | 14580 |

702

| | | | | | | |
|------------|------------|------------|------------|------------|-------------|-------|
| CCAAAAGGCA | CATGAACCCG | CATCCCAACT | TCCAGCATTC | CCTCAAATTC | CTCCGGAATC | 14640 |
| CTGTAACTAT | AGGGCTGGTC | CGTCTGCATC | AAGGGCACAT | CTACGATAAT | CTTAGCTAGG | 14700 |
| GCCATCTTCT | CACCTCCTCC | TTGTCACTAC | ATTCTTGCAA | TAGAAAAAAT | AAGATTGAGT | 14760 |
| CCCCCAACC | TTAAATTTT | TCACCATCTT | CTTTTCTTT | AGCAATTGTC | TCTTTGATTT | 14820 |
| TCTTTTCTTC | TTCTTCTTTG | CGGCGTTTTT | CTTCTTCGAT | ACGGCGACGC | ACTGCTTCAC | 14880 |
| GTTTTCCTTC | TGGATCTGGG | TGAATTGTAA | CGTTTCCTGA | TTCGATTTC | TCTAAAGCGC | 14940 |
| GAAGAGTTGA | TTTTTCAGAC | TTGAAACCTT | GAGTTGCTGG | GGCACCTGCT | TCCAATTCGT | 15000 |
| GGGCACGTTT | TGCTTCCAAG | ATTACGAGTG | AATATTTTGA | AGGAACCTTG | TCGAGCAAGG | 15060 |
| TATCAATAGA | GGGTTTAAAC | ATCATTTGCT | TGTACCTATT | TTCTAAATTT | TATCGGGTAG | 15120 |
| TTGGAGATTT | TGGTAACATC | TCCTGATAGT | GACCAATGAC | ACGATCCACA | CAGAAGTGTT | 15180 |
| CTGCTTCAAT | CACACATTTG | ACACGTTTCA | CAGCTAGGGG | TACCTGATCG | TTGACAATCG | 15240 |
| CATAATCATA | CTCACGCATG | AGGGCAATTT | CTTCCTTGGC | CTTTTCGATT | CGTTGGGCAA | 15300 |
| TCACTTCTGC | ACTATCTGTT | CCACGACCTA | CCAAGCGATC | TTGCAATTCA | TCCAATCTG | 15360 |
| GTGGTGTCAG | GAAGATAAAG | ACAGCATCTG | GAACCTTTTT | CTTGACCTGA | AGAGCACCCCT | 15420 |
| GAACCTCAAT | TTCAAGGAAA | ACATCGATTG | CCTTGTCCAA | GGTTTCATTG | ACATAGGTCA | 15480 |
| GAGGAGTTCC | ATAGTAGTTA | CCGACATATT | CTGCGTATTC | CAACATCTGT | CCTTGACGAA | 15540 |
| TCAGCTCTTC | AAATCTTTCA | CGAGTACGGA | AGAAATAGTC | AACACCGTCC | ACTTCTCCAG | 15600 |
| GACGTGTGTC | GCGTGTGCTC | ATCGATACAG | AATATTGAAA | TTGGTTTTCA | GAACCTCTCA | 15660 |
| AAATCTCTCT | TCTAACCGTT | CCTTTTCCAA | CCCCTGAAGG | ACCAGAAAAA | ACGATTAGTA | 15720 |
| AGCCTCGGTC | TGCCATTGTG | TCTCCTTTTA | GTCAATCTGT | GAAATAACAT | TTCTCTAGAA | 15780 |
| TAATGGCAAA | AAGCCAGATT | ATCCTTTACA | GTCTTTCTAT | CTAGTGTAAC | AAAAAAGCAG | 15840 |
| TAATTTTTCA | ACTGCTCTTT | CTTATTTATT | TAGCATAATC | TACTGCACGA | AGCTCGCGAA | 15900 |
| TCACGGTTAC | CTTGATATTT | CCTGGATAAT | CGAGATTGTT | TTCAATTTTC | TTACGAACCT | 15960 |
| TGTGAGCCAA | GATTGTGACT | TTGTGTCCTT | TGATTTTCC | TGGATTGACC | ATGATACGAA | 16020 |
| TTTCACGTCC | TGCTTGAAGG | GCAAAGCTAG | TTTGCACCTC | TTCAAAGCCG | TTAGCAATTT | 16080 |
| CTTCCAAATC | ATGGAGACGC | TTGATGTAGC | TTTCAAGAGA | CTCACTACGA | GCACCTGGAC | 16140 |
| GGGCTGCGCT | CAAGGCATCT | GCTGCAGCGA | CGATAACTGC | TATCACGCTC | TCAGCTTCAA | 16200 |
| CATCTCCGTG | GTGACTAGCA | ATCGTATTCA | CCACAACCTG | GGGTTCTTTG | TACTTACGGG | 16260 |
| CCAATTCCAT | ACCGATTTC | ACGTGGCTAC | CTTCAACCTC | ATGGTCAATG | GCTTTCCCGA | 16320 |
| TATCGTGAAG | GAATCCAGCA | CGACGGGCAA | GAGCCGCATT | TTCACCAAGT | TCGCTCGCCA | 16380 |

703

| | |
|---|-------|
| TGATACCAGC CAACTTAGCA ACCTCAATCG AATGGCGCAA AACATTTTGT CCATATGAAG | 16440 |
| TACGGAACTG CAAACGTCCC ATAATCTTCA TCAAGTCTGG ATGAAGGTTT GGCGCACCAA | 16500 |
| TTTCATAGGC AGCAGCCTCA CCGTATTAC GAATCTTATT GTCAATCTCT TGACGGTTTT | 16560 |
| TCTCAACCAA CTCTTCGATA CGAGCTGGAT GTATACGACC ATCTTTGAGC AACATTTCCA | 16620 |
| TAGTCATACG GGCAATCTCA CGACGAATCG GATCAAATCC TGACAAGGTC ACCACTTCTG | 16680 |
| GTGTATCGTC GATAATCACA TCGACCCCTG TCAAACCTTC AAAGGTACGA ATGTTACGAC | 16740 |
| CTTCACGACC AATAATGCGT CCCTTCATAG TATCGTCTGG CAGATGAACT GTTGAGTTTG | 16800 |
| TTGACTCCGC TACATATTCA CCAGCGATAC GTTGCATAGC TTGAACCAAG ATGTCCTTGG | 16860 |
| CCATTTTGTC AGAACGTTCC TTGACCTCTT GCTCAGCTTC GCGAATGCGA CTGGCAATCT | 16920 |
| CCCTGGTCAA GTTTTCCTCT GTCTGAGCCA AGATAATATC TCGTGCTTCT GCCTGAGACA | 16980 |
| GCGCACCAAT ACGCTCTAGT TCTGCTTCTT TTTGTCTTTC GACTTCCTCT AATTGCTCTT | 17040 |
| CACGCGCATC AAGGTTTTTC GCTCTATCAG AAATACTTTG TTCTTTTGT TCAAGTGT | 17100 |
| GTCTTTTACT CGTCAAATG TCGTCCTTAC GGTCAAGGCT AGTAGCTCTC TCTGTCAAAC | 17160 |
| GACTTTCGAT TTGTTGAGT TCTTGACGTT CTGATTTGAA TTCAGCGTCC ACTTCTTAC | 17220 |
| GGTATTTTCT GGCTTCTCT TTGGCCTCCA ATAGTGCTTC TTTTAAAGA GACTTGCTTT | 17280 |
| CACGTTTGGC TTCATTAACA AGTAAATCCG CTTCACGCTC AGCTTGTTCA CGTAAATTAG | 17340 |
| TTGCTTCTTG TTCAGCATTT AAAAGCATCA ACTCTGCAGC TTCCTGAGAT GATTTTCTCT | 17400 |
| TAGCTGAGAT GCTGACATAT CCAATGACTA AACCAATGAT GACGGCAAAA ACAGCAATCG | 17460 |
| CAAGCGACAT GATTTCCATG TTTTACCTC ATTTTATTGT TATCCGAAT GACATACATT | 17520 |
| CTTTTACATT CTACCATAAA AAAGTGATTT TCACAAACCT AAAATAGAAT ATGTTTTGAG | 17580 |
| GAATTTGGAA CACATTTACC AAAATAAACT TGTGTTTAG AAATAGTAGT TTAGTAGAGA | 17640 |
| CTTGAGAAAA AGCCTACCTT TCAATAGACT TAGTAATGAT CTTTAAAGGA CAAGAAAGCC | 17700 |
| ACGCTATCTC CATCCATCAT ATAAATCAAG CGATTTTCTG CATCAATACG CCGTGACCAG | 17760 |
| GCTCCTTGGT AATCATATTT GAGTGGTTCT GGTTTACCTA TTCCTGTAAA GGGATCACGT | 17820 |
| TGAATATCCT TGATTAGTTT ATTGATTCTT TTTAACGTTT TCTTATCCTG ATTTTGCCAG | 17880 |
| TAGCAATAAT CTGCCCAGGC ATCTTCTGTA AACTTGAGCA GCATTTCTTA CTCCTCAATA | 17940 |
| ACATGGACCT GAGTACTTCC AGCACGAACT TGAGCCATTC CTCGCAAAAC CTTATCAGAA | 18000 |
| AGTTCCTTAT TTTGAGCAAT TCTCAGGGTT TCTTGGATAC TATCCCACTC ACTCTTTGAA | 18060 |
| AGGACTACAA TGTCTCATC TGGATTTTTA TTGACCACCG TCAAAGGCTC AAATTCATCA | 18120 |

704

| | |
|---|-------|
| TTTACCTTCT TCATGTAGTC CTTTAAATGA TTTCGGAATG TTGAGTAAAG GACTGCTTCC | 18180 |
| ATAACCATAC CTCGTTTTC TAGCTCTTTTCCA CTATTATACA CGAAAAGAAA GAAATTGTCA | 18240 |
| GGAACCTGTA CAAGATTTTC TTTTCTATCT ATTTATACTC AATGAAAATC AAAGAGCAAA | 18300 |
| CTAGGAAACT AGCCGCAGGC TGTACTTGAG TACGGCAAGG CGACGTTGAC GCGATTTGAA | 18360 |
| TTTGATTTTC GAAGAGTATT ATTCGTAAAA AATCTCAAAA AGCCTACCTT TCGGTAGACT | 18420 |
| TAGTTTGTTC CTATTC | 18436 |

(2) INFORMATION FOR SEQ ID NO: 88:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7001 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 88:

| | |
|---|------|
| ACGTAGAAAA ACTATTTCTA TCACAGATAA TATTCCTAT GTTGTGGAG GTATTGAAAT | 60 |
| AAACGTCCTA GGTATCTTTC TCAGTCTATG TGACTTACAA GGGAAACTC TTTTCGAGAC | 120 |
| AGAAATTTTG AATGAAGATT ATCCTATTTT AGAAATCAAT TCCACCATTA CCAATATGAT | 180 |
| AAAAACAGCT ATAGAGTACG TCCCTTTGGA AACAAAATTA CTGGGATTG GCTTATCAAT | 240 |
| ACCTGGACAT TATAACAAAG ACTCCGGAAG TATCATTACA AACACCCCA TATGGGAATC | 300 |
| TTTTAATTTA TTAATGTAA TTAAGATT CAATTTTCCT TTTATTGTAA AAAATAATAT | 360 |
| CGATTGTATG GCTATAGGAC AATACCTTTT TAATCCACAC AATACCCCG ATAACCTTAT | 420 |
| TTTCCTACAC GCTGGATTAG GTATTACAC TTCCTTTTC ACAAAGAAA AAATAGGAGC | 480 |
| CTCTAAAAAT CCTTATATCG GAGAAATTGG ACACACCATT GTCGAATTGA ATGGGCAATA | 540 |
| TTGTGAATGC GGAAAAAAG GTTGTTTACA AACATATATT TCGGATGCTT GGTTAATCAA | 600 |
| ACACGCCCAA TTATTATTTA AAAATTCCCA ACTAACTGTA CTAAAAAGCC TTGTAAAGAC | 660 |
| TGAAAAAGAC ATTCATTTAG ACACCTTTT AACGGCTTAT AATTTAGGCG ACTCCGCTTT | 720 |
| ACGTCAACAA ATTGATAAAG GAGTCAATTT ATTAGCCACT TCTATTGCAA ATCTCCTCCT | 780 |
| CATCAATCCT GCTGATAAAA TCTATATCAA CAGTCAATTG CTTAATTATC AACCTTTCAC | 840 |
| TCATGAAGTC AGGGATAAAA TCCAAGACCA GCTCCACTTC GTTCCCTTTA CTCGTAATAT | 900 |
| AGAAATTGAA ATTTTACCTT ACAACAAACA TCGTGGAAGT ATAGGAGCTT GTGCATTAGC | 960 |
| TATCGTCGCT TTTTTCATAG AACATAGCAA TGTATTACAA GATATTATTT CACCTTAATA | 1020 |
| TATTAGAAAT CTATAGACCT GTTTAAATCA ACTATAACCT GTAGTAGATA TCTCGTATTT | 1080 |

705

| | |
|--|------|
| AGACAATATG AAAACAAGAC GACTTCCATA TAGGAAACCG CCTTCTCGCT ATGTTGAGTG | 1140 |
| ATTTATATTA AAATAACTTT TCTTCTAGCT GCATTTTATT ATTATAAAAA CATTCATCAT | 1200 |
| AACCCCAGA ACTTAAATA CAATTTTAT TCAAGATACA TACTCCTAGA ATAACTTTA | 1260 |
| TATGAAATTC TCATTTTGT TTTTACAATT CTCCTTAGTT AAATCTTGT TAATATATGT | 1320 |
| TTTACATATA GTATTTAGCG CCACATAGTA CTGAACTCTC TCCAAAAACG GTTATTCCCTC | 1380 |
| TTTGAATAGG GCGTTATCAC AAGAAAAGCA TCTCCACGTT TCAACTTCAT ATGGCTCAAA | 1440 |
| AACAATCAAT TGATGCTAAA ACCTGTACCT AGATGTTTCG GTTCATAAAA CCATGAACT | 1500 |
| GTAAAAGTGG ATGAAATTGA TAGCGATAGT CAAATCAAGA GGCATCATAA CTCTAAAAAG | 1560 |
| TCACATATA TAAGTTCATC CTCGGAAGAA TATCATTCTA ATTGTTGAAA TGCCTACATG | 1620 |
| AAAAGAAACG TCAAATGCTC ATGAAACAAC GAATACAGGT ATCAAACTA TGACAAAACA | 1680 |
| AATCCCTAAA TTTACTAAAG ACACTGCTCA ACTTTACACC TGTAATGGT TGTGTATATA | 1740 |
| TAAAGTTACA AAGATGTACG ACCACACTGT TGTAAATCAT AGTGTTCGCG AATATATTAC | 1800 |
| TGATAGCATT TCTACAAATA CAAGTAAAGA GAGCGGATGA GATTCAAACG AAATATGTCA | 1860 |
| GTGCTTTGGC ATTCTAGGCC TTCATATCAT TTAAGAATT CTATAGACAA AATTTTTC | 1920 |
| AATACAGACA CTCGTAACAA CTGCTTCATT TTTCTACCA CATATTTAGG AACAGGATAA | 1980 |
| GATACAAGAG TATTAATCCA TAGCTCAGTT CTATACCAAT CTAAGACAAA TAAGCTAAAA | 2040 |
| AAACGATTGA TAATAAGCAA ATAGATTCCA AATTTTCTCT ATCTGCTCAT TTTAATAAAC | 2100 |
| AATACTAGTG TAACTATCCT TCCAGTCAGA AGCTTGTCAG ATCACACCGA AAATCTTCT | 2160 |
| AAAATTTATC TCGTTAGGCA ATCAAGCAAA AACTCGACGA TAGTACAAAC ATTATCATAC | 2220 |
| AGGATTGACT TCCTAAATTA TATACTTTAG TAAGGTTTTC GGATAAGAAA AAAGGTTTCT | 2280 |
| TTTACATTTT TAAACATTCT TTTCTAAGAT GAAAAACAGA ATTTTTCGAT TGTGATTTAA | 2340 |
| AGCAACAAGA AGATTTTCAG TATCATCCTA TAGATACGAG CTAATTAAGA AAACTACAT | 2400 |
| TTTTGAATAT AAACATAAT AATATAAACT AAATTTTATA GGAGGAAGAC AATGGATTGG | 2460 |
| TACGATTATA TGATACAGGC ATCCAAACAA TCACAATCA ACGCAAGCCA TTGGTTTCGC | 2520 |
| TATTTGCGAA AAGTTATTTT TGAAGACTAT TCTTATTTAA CAAACCAAGA TGTAGAAAAG | 2580 |
| TTGCTAGACT CCAAAGAACT AACCCGTTTT CAAAAAATTA GCTTGAAGTA TGCCTTTCAA | 2640 |
| GAGCATACTC CAACTCATAA ATATGTGATT TCATTAAATA AACCTGCTAA GTTAACCAAT | 2700 |
| GTTCAAAAAT TGATGGAGAA ATACAAACAT GGATAAAATG AAACCGGTCT TCCAAGCCCT | 2760 |
| AAATAAGGAA TTAATTCAGG AAAATCTGAC TTTAACAATT ATCTGTGTCG GTGGTTATGT | 2820 |

706

| | |
|--|------|
| CTTAGAATAT CATGGTTTAC GTGCCACACA AGATGTTGAT GCTTTTATGG CTCTATAATA | 2880 |
| TTTGTAGTGG GTAAATCCCC TATGGATATT ATGGAGCCTA TTTTGTGTGA GAAAAAAGT | 2940 |
| CCCATATGAC CTATAATGAA AAGCGACAAA ACAACTCAT T AGAAAGAATC ATATGGAACA | 3000 |
| ATTACATTTT ATCACAAAAT TACTAGACAT TAAAGACCCT AATATCCAGA TTTTAGACAT | 3060 |
| CATCAATAAG GATACACACA AGGAAATCAT CGCCAACTG GACTACGACG CCCCATCTTG | 3120 |
| CCCTGAGTGC GGAAACCAAT TGAAGAAATA TGACTTTCAA AAACCGTCTA AGATCCCTTA | 3180 |
| CCTCGAAACA ACTGGTATGC CTTCTAGAAT TCTCCTTAGA AAACGCCGTT TCAAGTGCTA | 3240 |
| TCACTGTTCA AAAATGATGG TCGCTGAAAC TTCTATCGTC AAGAAGAATC ATCAAATTC | 3300 |
| TCGTATTATC AACCAAAAAA TTGCGCAAAA GTTGATTGAG AAGATTCTA TGACCGATAT | 3360 |
| TGCTCATCAG CTGGCCATTT CAACTTCAAC TGTCATTGCG AAGCTCAATG ATTCTCACTT | 3420 |
| TGAGCATGAT TTTTCGCGTC TTCCTGAGAT TATGTCCTGG GACGTTGAAA CAGTCCGGGG | 3480 |
| AGTGACTGTT TCAATCGGGA GATGGAGATG AGCTTTATTG CGCAAGATTT TGAAAAGCTC | 3540 |
| GATATCATCA CTGTTCTTGA AGGTAGAACA CAAGCTGTCA TCCGAGATCA CTTTCTTAAA | 3600 |
| TATGATAGAG CCGTCCGATG TCGCGTCAAA ATTATTACTA TGGATATGTT TAGTCCTTAC | 3660 |
| TATGACTTAG CTAGACAACT TTTCCCGTGT GCTAAAATCG TTCTTGATCG CTTTCACATT | 3720 |
| GTACAACATC TTAGCCGTGC TATGAGTCGT GTGCGTGTCC AAATCATGAA TCAGTTTCAT | 3780 |
| CGAAAATCCC ATGAATACAA GGCTATCAAG CGCTACTGGA AACTCATTCA ACAGGATAGC | 3840 |
| CGTAACTCA GCGATAAACA TTTTATCGC CCTACTTTTC GTATGCATTT AACCAATAAA | 3900 |
| GAGATTTTAG ACAAGCTTTT GAGCTATTCA CAAGACTTGA AACATCACTA TCAGCTCTAT | 3960 |
| CAACTCTTGC TGTTTCACCT TCAGAATAAG GAACCGGAGA AATTTTTCGA ACTTATCGAG | 4020 |
| GACAATCTTA AGCAGGTTCA TCCTATTTTT CAGACTGTCT TTAAAACCTT CCTCAAAGAT | 4080 |
| AAAGAAAAGG TTATCAACGC CCTTCAACTA CACTATTCTA ATGCCAACT GGAAGCGACC | 4140 |
| AATAATCTCA TCAAACCTTAT CAAGCGCAAT GCCTTTGGTT TTCGAACTT TGAAAACCTC | 4200 |
| AAAAACGGA TTTTATCGC TCTGAATATC AAAAAGAAA GGACAAAAT TGTCTTTCT | 4260 |
| CGAGCTTAGC TTTTTCCTT CCACTACAG TTGACAAAGA GCCGAAAAA GGAACAGCCT | 4320 |
| TAGCTTTCCT TTCATTCTT TTTATTTCCC TCGTAGTAAA CGTGCTAGCT TCCACAAAAC | 4380 |
| AAACAGGATT CCCAGAAATG CCAGTACCAC TAGCCACGG TACAACCAT T GAGAGGTTGC | 4440 |
| AACACGCGAT ACAGATTGTC CTTCTTTCGT AAAAGCAACC CTCGCAACTG CAGCTGTTTG | 4500 |
| TGGATCTGAT TTTTGATAAA CAGCGACTCG TTCAAAATTC ACTAATAAGC GTTTATTAAA | 4560 |
| GGTAGGAATC GGATCGCAGG TTATCAAGGT CATGATATTT TTAGAGCTAA CCGATTCTAA | 4620 |

707

| | |
|--|------|
| TTTTTCCCAT TCCGACGGTA AAATAATCTC TGTGTCCATC ATCTGATATT CTACAATTTC | 4680 |
| CTGGCCATTA TCATAATAAA GAGCATCTCC AACTTTTAGC TGATCCAAAT GGCGGAAAAA | 4740 |
| GACATGGCTT GGCTCTGCAC GGTGCCCAGC AATCACTGAG CGAATCCCTG TACCATCCAG | 4800 |
| AGGCAGCGGT GTACCATCCA CATGAGCCAA GCCCATCCCT AAATGATGAT AATCTGCTCC | 4860 |
| CAAATAAACC GGCTCCATGA TTTCCAACT TGAATAGAC AAGTAACCAT AGACTGCATC | 4920 |
| AGGGTCGTCA GACACTTGGT AATTGACCTC ATATCCCTCC GCCAAAAAG GATCTACAAT | 4980 |
| GCGATTTTGC GAAGCCAAGC GTTGATTGTA GCGAGAGAA TGGTTCTGTT GTTCTTGTA | 5040 |
| CATTTTCAGTT GTCATGGATT TCACAAATGT AGCATGACCT TTCACCTGTC CAAGAGACTG | 5100 |
| CAACACCATC TGTCCAAAAC AATAAATAGG AATCAAACAG GCTACCAACA TCAACAAGTA | 5160 |
| TCCAATAAG GCTCGTAGTT TAGTCCTTGA CATGACGCC CTCCAATTGC TTTTCTAGTC | 5220 |
| CTTTGACAAT CCGTCGATTA CGATACACGC GATACAGCAA GAGAAGGATG ACCGCCATCG | 5280 |
| CTCCTAGTAA TAACCACAAC CAGAATTGCC CACGCTCTCT CACCGCTCGA TTCCGCTCTG | 5340 |
| CAATTGGTGC CGTATACGGA ATCCGCTTCC CACGTACCAA CAGACGATGA CTGTTAATCA | 5400 |
| TATACGGTGT ACAAGTCAAC AAGGTGCGAT AATCTTCCCC ATGTTGAATC AAGACAGGCT | 5460 |
| CAAAGTCATT CGGCTCCACC GTCATATCT GATCCACTTG GTAGGCCAAC ACCTGATCTA | 5520 |
| AAACGTGAAG ATAAAAGATA TCCCCTTTT TCATCTTATC CAATTGACTG AACAAATCTG | 5580 |
| CCGTTGGCAA TCCTCTGTGA GCAGTGATCA CTGTATGGGT ATTTTCACCT CCAACAGGCA | 5640 |
| GCGAAGCCCC TTCTAACAGC CCTGCCCCCT TCTGAAGAAT GTCCTCACTC GTTCCGACAT | 5700 |
| ACATCGGAAT TTCTGATCA ATCGCAGGAA TTTCCACATA GCCAATCCGC TCATGGACCT | 5760 |
| TTAGCATATT GGCATATTCT GAGACGCCTT TCTTTTCTC TTGCTCTGTA AAAGGATCAA | 5820 |
| GAATTCAGA TGGTTTCAAG GTCGCATTGA AGGCTTGAGC CAAGCGCCAA CGCTCCTCAA | 5880 |
| GTTCCTGCCTT ATCCATCTGG GAAACCGTCT CATCAAATC TTAAATAACC TCGTTTGACT | 5940 |
| CAATACGATA ATAATAACGA GACACCAATG GATATATCGC AACGGCGAAT CCTACTAAGA | 6000 |
| AAATCAGAAG AAGGATCAGC GGATGTTTCT TCTTTTTTGT GCCTTTTTTTT CGTGAACGTC | 6060 |
| TACTGTTGTC CATCCTCCAC CTTCCTTCC TTCCTTGCTG CTTCAGCGC CTTCAAAGCC | 6120 |
| TTTTCCGGTT GTTTTTTCTT CTGCGCAAG CGTCGAATAA TCCATAAAAG AATCACAATC | 6180 |
| AAACCAACTG CCACATAAAA CAGGTAGCGA TAGAGATGAC TGAGTTTGTT TGCTGCAATA | 6240 |
| AATTCTTCCT CAACCTCTGC TACGTACGGT ATCCGATGCC CCCGAACCAA TAGACGATGG | 6300 |
| GTATTGATCA TGTATGGCGT ACAAGTCAGC AAGGTCACAT AATCATGACC TGGTACAATC | 6360 |

708

| | |
|--|------|
| AATAAATCAT CAAAGTTTCGT CGGCTCAATC ACCTTTACTT GATCCACTTG ATAGGCCATC | 6420 |
| ACTTCCTTGA TATTGTGCAC ATAAAACTTA TCCCAACTT TAAGTTGGT CAAATCCGTA | 6480 |
| AACATCTTAG CTGTTGGCAA ACCTGTATGT GCCGTAATCA CCGCATGGGT CGAATTGCCT | 6540 |
| CCGATCGGCA GAGAAGTTCC CTCTAGATGC CCAGCCCCTT GCTGCAATAC CTCTTCAGCA | 6600 |
| GTACCAGCAT AAACCGGCAA ATCCACGTCA ATAACGGGGA TTTCCACATG CCCCATCCGC | 6660 |
| TCATGGATTT CTACATACG TGCATACTCT GCTCGCCCTT TTTTCTTCAT TTCTTCCGAC | 6720 |
| CAAGGATCGC CACTCACTAC ATTATTCAA GAGTCATTGA AGGCTTGTGC CAATTTTCATT | 6780 |
| CGTTCATCAA TGTGAGCCTC ATCCAACGTT GCTTTTTCCT TATCAAAGTC AGCAATTGT | 6840 |
| TGATTTGATT CCACTCGATA ATACAAGCGA GACACCAGCG GATACGCCAT TACCGCCATT | 6900 |
| CCAATGAAAA ATACCACTCC TAATAGGAGA TTATTTTCGTT TTTGCTTTT TGTTTTACC | 6960 |
| ATTTTATCA GCATCCCTTT ATCTTCAAAC TTCAGGGTAT C | 7001 |

(2) INFORMATION FOR SEQ ID NO: 89:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10411 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 89:

| | |
|---|-----|
| GAGGGAGCTT AAGAAGTTAC CACCGTCCTC TAGCGCCTTA TCCGCATCAA AGTTAAGGTT | 60 |
| GATATTTTGA AAAGTGTGCG CAGCTTGTGA TACGATGCTT TGTTTAAGGT CATTTAGGGT | 120 |
| TTTAGTGAAA TCTGCATTGC TGAGGATATC ACTCTTTGAG AGATTCAAGG CAAAATTGAT | 180 |
| GATGATATTG ATCTGGTTTC CTGTTATGAC CTGATCAAGT TTGTAATTTT TTAAGGTATC | 240 |
| TTCAACAATC TTGCGGATAT CTTCTTCTGT CAGATTTCCT TTACTTTCTT TAGCTTTGGC | 300 |
| GAGTCCTGAC TTGATATCAG CTAGGGCAAC GTTTAATTTA TTAGCATCAT AGCCTGATTT | 360 |
| GTCCTTGTTC TCAGCATTGA TATCTGACAA AGCTTTTAGC TCTTCTTGAG CCAAATCTTT | 420 |
| ATTAGCTTGT GGCACCTTGG CTCCATTAGC CTCTAGCGAA TAGTAAATCC CTGCTAAAGC | 480 |
| ACTTTCTCCT GTAAGTGAA TAGGGGCTGC TACAGTGATT TTGGCATGTT CCATACCCAG | 540 |
| CGTTACTGCT GCGTTTCGGT ACATATCCTG AGTCACCTTA GTGATATTTT CTGGTGTTTC | 600 |
| AATCTTGACC TCAAGTGGCG ATTTGTCACC TAGCTTTTGA ATCTTGGCTG ATGAATACAA | 660 |
| CTGTAAGCTA GAGTCATTGG CCACATTCAT GATTTTAGAA TAAACATCAG GTGTCATGGT | 720 |
| CTTGAGTTCT TTGGTATCTG TTGAGGCATT GTAGCCAGT TTTTAAAGAG TTTGATTTTT | 780 |

709

| | |
|--|------|
| TTGGTCTTCA GATAGGGAGG AACCTAGGAC ATATTCAGGT TGGACATAGG TTTCATCGAT | 840 |
| AACCTTTTGA ACATCTGTTG CTGCATGGAC GCTATTCATA GCTGTACTG CCCACAAGAT | 900 |
| CGCAGCGCTA GTCAGAAAGA GTTCTTTCT CATAGGGAAT TTCCTCCTTT ACTTCTTTAG | 960 |
| AGTAATATAT CTATCTTAAA GAAAACCTAT AACAAAAACA CCTGGTCTAG CCAGATGTTG | 1020 |
| AAAAGAGAGT GAAACATTTG ATGATGTAAA GGTAAAGTCG TACCTGTCTA GAATAATAAT | 1080 |
| AGTTTCCTCC ATTTACATAG AGTTCAGCAC CGTGAAAAAT GGAAATGGGG TGAATATAAC | 1140 |
| TATAAGTCTT TCCAGTCCTA TTACCAAGCA AGGGGGCAAC AGTCTCACGA GAGTACTGTT | 1200 |
| TGGCTAGAGC CAGGGTATTT TCCTTGCCAT TTTGGGCGAT AAAATCGATA TAGGCAGGTC | 1260 |
| CAAAATTATA GGCTTGAACA GCTGTCCAGA TATCTACCCC CTTCTTCTGC GCCAGATAGA | 1320 |
| GATTGCCTGT CAGAGTTTGA ATGCCTTGCC GAATGCTAGA GGCATTATCA TTGATGGTGT | 1380 |
| TGGTGAAC ACTTGCAGAC TCACTAGACT GCATAACATC GCCTTCTTTT CCTTTTGT | 1440 |
| CAGTATAAAT CATAGCAAGC ACAAGCTCTT CGTTTGCTGG GGTGTCTTGT TCACTCAATA | 1500 |
| TTTCTCGCAC CATGGGTGA TAGGTCATGA CTGTGTTGAC ATCTTGATGA ACGCGGTAAG | 1560 |
| CTTTATAGCC AGCAAAAAGG AAGACTGCTA GTACAAGCAC TCTTCGAATT CGTTTAAACA | 1620 |
| TTATTTACTT TGGATATCCT CGATATTTTT GATTAAGATA GAGTAGGTTC CATTTTCGTT | 1680 |
| TTGGATAAAC TCAACAGACT CGGCGTCTTG ATAGACGTTA TTGGGAACGA TGAGCTCAAT | 1740 |
| TCCATTTGAT AAGGAGAGTT TTTGGTTTTT AAATTTCTTT AATTGGCGAC TGGCATCAAT | 1800 |
| TTCATCAAAT TGAACAGGTT CTGGTACGGC TTCTTTGACT TGGTCAATAA AGCTCAAACG | 1860 |
| AGCCGTCAGA TTGTTGTCAA AAAGGTCATT AGCCAATTTT TCAGGTGACA ATTCATGTCT | 1920 |
| TTCTTCTAGG TTGTTGAAAA TAGCTGATTT GACCTTGGAT TGAAATTGAA AATCATCTGT | 1980 |
| GTTAAAAGAT TTAGCAATTC TCTGGGCTGT TTTTCCAGT TCCTTGATAG ATTTTITAGG | 2040 |
| AGAAATCTTA GGAGCGACAG CAAGAAGATT ATCTGAAAA TAGTTCAAAA AAGTCCCGTT | 2100 |
| GTACTTGATT CGTTTTTCAA TCAGGTGATA CTTGCTACTC TGAAGATTGA CCACCAAGGC | 2160 |
| CTCATCAGCT CCTGTTCCAA ATCCAGGCAG GTTATTCTGA GTTAGCTTGA TTGGATTATC | 2220 |
| AACCTCTCCT CCGAGGTGGG TCAAGGTCTC CCGCAGGGCA ATTCGCAAGA AAGCGAAATG | 2280 |
| TTCTACACCT TCTTTAGAAA ATTGCACAAA AATCAAGTCA TTGGTCTTGA GATTTTCAGA | 2340 |
| AATGCTAAAC TCCTCTTTCC AGAGATTAGC CAGCGTACT GATGTCTCCA ACAAATCGTC | 2400 |
| TGTAATATGA TTGAAGAAGG GATTTTCTTC TTCGAAAAATC CCAGTCTTGG CTTCATCTGA | 2460 |
| ATACACATGT TCAATTTTTT TACGCAGGTA TTCTTCGATT TTTGGAGTAA TATTGAGAAA | 2520 |

710
CTTATCTGCT AAGAACAGTT CGGTATCATC CGGACTGAAC TGGTGAATAA TGGCTTTCTT 2580
AATATAAATG TCCATAAAAG TTTTAGTCCT CGTATAATGG GAAGGCATCT GTCAATTCTT 2640
TGACTGCACT TCTCACTTCT TCTAATACAG CCTCATTTTC TGAATTCTTA AGGGTTTAA 2700
TGATGAGTTC AGCCACTTTG CGACTTTCTT CTTCACCAA TCCACGTGCA GTAATGGCTG 2760
CTGCTCCGAT ACGAATCCCA CTTGTCTTGA ATGGTGACAA GCTTTCGTAA GGGATTGAGT 2820
TTTTATTAA GGTAATATTG ACTTCATCCA ACAAGTTTG AGCAACTTTG CCGTTTCTA 2880
CAACTTTAGT CACATCAACA AGGAAGAGAT GGTTTTCAGT TCCACCTGAA ATAATACGGA 2940
AATCAGGGTC TTGCAAGAAG ACATCTGCCA TAGCCTTGCT GTTCTTAATT ACATTGGCAG 3000
CATATTCCTT GAAGGCTGGA TCCAAAACCT CTTTGAAGGA AACTGCCTTA GCCGCCACAA 3060
CATGCTCTAA AGGACCGCCC TGAATACCTG GGAAAATAGC TGAATTGATT TTTTTCAGAA 3120
GTTCTTCGTC ATTGGTCAAA ATCAAACAC CACGAGGTCC ACGAAGGGTT TTGTGGGTCG 3180
TTGTTGTTGT GATATGAGCG TATGGAATG GGCTTGATG AAGGCCAGCC GCAACCAAGC 3240
CAGCGATATG GGCCATGTCC ACCATGAGCT TCGCACCGAC AGCATCTGCG ATTTACGGA 3300
ATTTTGAAAA ATCGATAATT TGAGAATAGG CTGAAGCACC AGCTACAATC AGTTTGGTT 3360
TTACTTCTTG GGCTTGTTC AAGATAGCAT CAAAGTCTAA GAGTTCCGTT TTAGGATCAA 3420
CACTATAAGA AACAAAGTTG TAGGTTTGAC CAGAGAAGCT AACAGGAGCC CCATGAGTCA 3480
AATGACCACC TGATGCCAAA TCCATTCCCA TAACCGTATC ACCTGGCTCA ATCAAGGACA 3540
TGTAAGCCGC ACAGTTAGCT TGGCTTCCTG AATGTGGTTG AACATTGGCA AATTAGCAC 3600
CGAAAATTTC TTTTGCGCGT TCAATAGCAA GAGTCTCTAC AACGTCTACT ACATCAGTTC 3660
CACCATAATA ACGGCGTCCT GGGTAACCCT CGGCATATTT ATTTGTCAAG ATAGACCCTT 3720
GAGCTGCCAT AACAGCCTTG GAACTACGT TTTCCGAAGC AATTAACCTG ATATTATTTT 3780
GTTGGCGTTC TTCTTCTTG GCAATAGCAT TCCAGAGATC AGCATCATAT GCTTTAAAT 3840
CATCTTTGTC AAAAATCATA GGTCTTCTCC TTTATTGTGT GACTAGTCCA TTAGTTTGAT 3900
TTTACAATAA GAAAATCAAA CTAACAGATG CGAATAAACC GTTCTGTCAT TTTATCACAA 3960
GTATAGCCAA CTTTTTCATA AAATGCATGA GCACCCAGAC GATGATTGGC AGAATTTAAG 4020
CGGATAAACC CATAACCACA TCTTTTGCT TCTTCTCCA ACCCTGTAG TAACTTTTA 4080
CCAATACCTT GACCTTGCGC TTGAGGTGAA ACTGCTAAAG CTAAGATATT AAATCCTGCT 4140
TTGGAATAGA GTGATTGCTA AACTTCAGCG TGGACATATC CAAGTAAGAC ATGATTAGCT 4200
GCATCCTCAT AGCCAAGTAG GAAATGATGG GAATCCTGAG ACAGTCTAGC TAGTTGGCTA 4260
GCCGTTTCCT CTGGAATAAA AGTATAACCC AAAGCCTCTT GGTGATGTC ACATATAGCT 4320

711

| | | |
|--|------------------------|------|
| TTCACATCAG TTTCTCTTAA ATCTCTTAGC ATCTCATTC | TCCTCAAAAG AAATCTTTGG | 4380 |
| CAACCGAGCA AGAATATCTT CTCGCTTAAT GGCCCTTGA | CGTAAGATTT TCACCTTGTC | 4440 |
| TCCCACAAA TTCAAATAG TTGAATCCTG TCCAGTTAGA | AAAGCATCGT CTTCCAGACC | 4500 |
| CAGAACCTCT TGGTCAAAAT CCTCTAGAAT TTGATTAAAG | GTCACTCCAC TCGCCTGACC | 4560 |
| TGAGATATTG GCAGACGGCC CAATCAAGGG ACCTGTCTCT | CGAATCAAAT CAAGGGTAAT | 4620 |
| GGGATGACTA GGCATCCGAA ATCCAACAGT TGCAAGGCCA | GAATTGACCC AATAGGGAAC | 4680 |
| TCGGTCATTA GCTTCGAGAA TAATGGTCAA GGGACCTGGT | AAAAAGATCT CTACAAGTTT | 4740 |
| TTGAAGATAA GTTGGCTGAT TCTTTGAAAA GTACAAGATG | TCCTCTAAAG AGGCAACATT | 4800 |
| GAGATTGAGC GCCTTGCTCT TACGTGACG TTTAAGCTGG | TAAACATGGT CAACTGCTTT | 4860 |
| TTCTGTCTAGC GCCTTAGCAA AGAGACCGTA AACTGTCTCT | GTAGGCAAAA CGACAGCTCC | 4920 |
| ACCATTTTCC AACTCTTGTC TAATCCTGTC CATCATCAAC | GACAACCATC CTATCTTGAC | 4980 |
| CAAATGGTC CTGAGTGTT CGTACTCGCT TTTCAGGAAG | ATGTTTCCTA AAAAGTTCAG | 5040 |
| GAACACTTTG ACCTTGCTTG TATCCAATT CAAGGTAAAT | CTTACCACCA TCTTTGAGAT | 5100 |
| AGTCTTTTGC ATCTTCCGCA ATTCTACGGT AAATAGCTAG | GCCATCCTCA TCTGCAAAGA | 5160 |
| GAGCTAGATG AGGCTCCGAA TACAAGACAT TCAAGCCTAC | CTCTGACTCA TCTTCACGAG | 5220 |
| AGATATAGGG TGGATTGGAA ACAATTATAT CATATTTTTC | AGAAATTTCT GTAAAACAGT | 5280 |
| CAGATTTTTT TAAAAATATT TGAAGATTTT GATTTTTCAG | ATTTTCGCTA GCTACATCTA | 5340 |
| AAGCATCTTG GGAAATATCT GCTGCCGTCA CTGACCAATC | TGGTCTGTTT TTTGCTAGAG | 5400 |
| CGAGAGCAAT AGCTCCACTA CCTGTTCCGA TATCTAGGAC | CATAAGATTT TTCACAGGAT | 5460 |
| TTTCAGCCAG GATAAGCTCC ACCAACTCCT CTGTTTCTGG | ACGAGGAATC AAAACCCGTT | 5520 |
| CATCCACCTT TAAATGCATT CCATAAAAAT TGCCTGTCC | AATGATGTAC TGAGCTGGCT | 5580 |
| TGTGAGCTGC TAGTTGCTGG TAAATATCTT CTACAAATTG | TTTTTCTTCC TCTGTGTCA | 5640 |
| CCTCCTGCTG GAGGGCAAAA ATAAAGTCTG TAAAAGATAG | ATTTTTCAGA CTACGATAGA | 5700 |
| CAAAGAGAG GCTTTCGGCT TCCTCTCCTT GTCTTATCAA | CTCTTCTTCA AAATTGAAA | 5760 |
| ATAATTGAGC TAATTTCATT ATTTGTTTAA TTCTTCTAGT | TTTTGTGTTT GGTCAATAAG | 5820 |
| CACCAAGGCA TCCACAACCT CGTCCAATTT ACCAGACAAA | ATCGTATCTA GTTTTGGAG | 5880 |
| GGTCAAGCCG ATACGGTGGT CTGTGACACG GTTTTGTGGG | AAGTTATAAG TTCGGATCCG | 5940 |
| TTCTGAACGG TCACCAGTAC CGATTGTCGA CTTACGCTCA | GCGTCCTGCT CATCTTGAGC | 6000 |
| AATCTGAGCA AAGTGGTCAG CAACACGGGC ACGGATGATT | TTTCATGGCCT TCTCACGGTT | 6060 |

712

| | |
|---|------|
| CTTCTGCTGG GTACGTTCTT CCTGCATCTC AACCTTGATA TTGGTTGGCA AGTGAACGAT | 6120 |
| ACGAACGGCA GTCGCAACCT TATTGACGTT CTGTCCACCA GCACCAGAGG CGTGATAGAT | 6180 |
| GTCGACACGA AGGTCTTTTG GATCAATGTC GTATTCAACC TCTTCAACTT CTGGCATAAC | 6240 |
| AAGAACTGTC GCTGTGGAAG TATGAACACG GCCTTGGCTT TCTGTCACAG GAACACGTTG | 6300 |
| CACACGGTGG GCACCTGATT CATACTTAAG CTTAGAGTAT ACAGACTGAC CTGAAACCAT | 6360 |
| AGCAACCACT TCTTTAAAC CACCGACACC ATTCATAGAG GCTTCCATGA CTTCAAAGCG | 6420 |
| CCAACCTTGG GCTTCCGCAT ACTTTGGTA CATAGTTAGC AAATCTCCAG CGAAAAGTGC | 6480 |
| CGCTTCGTCT CCACCAGCTG CTCCACGGAT TTCAAGGATG ATATTCTTGT CATCGTTTGG | 6540 |
| ATCCTTTGGA AGGAGCAAAA TTTTCAGTTT TTCTTCATAT TCTTCTTTT CAGCCTTGGC | 6600 |
| ATCTTTGAGT TCTTGCTTGG CCAATCTTC CAAGTCCGCA TCTCCGCTG ATTCCTTAAT | 6660 |
| CATCTCTTCG GCATCGACGA TATTTTGAAG GACTTGTTTA TACTCACGGT AGGCTATTAC | 6720 |
| GGTGTACGA TTGGAAGCTT CTTCTTTTGA AAGCTCCATA AAACGCTTGG TGTCTGAAAC | 6780 |
| GACATCAGGG TCACTCAGCA ATTCTCCTAA TTCTTCATAA CGGTCTTCTA CAACTTGTAG | 6840 |
| TTGATCATAG ATGTTCAATT TTTCTCCTTA TTTCTCAATT GTTAAATCAT AGATTGCTAC | 6900 |
| TACTTCATTC TCGGATATTT CCCCAGTTTC TTTAAATCCA TAACTGAGGT AACAAAATCT | 6960 |
| TGCCTGTTCA TTTTCTGGTT CATAAGACAA CCAAAGTTTA TTGCTTAAAC CTGCTGGCGC | 7020 |
| TGTTGGAACA TAGTCTAGTA CTTTATCCAT AATTGGTTTA AAATATCCTT GATTTTGAAA | 7080 |
| ATTCTTATCA ATCATAAAAC GAAATAGTAA ATAATTTCCA CTACTAATTC CGATCTTTTT | 7140 |
| ATCATAAGCT ATCATCACAA AACCTATAAT TGCATCATTA TCATAAACTG CCAATGGAGC | 7200 |
| TACAAAATCT CCATTTTTAG TGTAGACGTA TGCTTCAGCT AAATAATTG CGTTGGTTGC | 7260 |
| AATGAATTGT TTTTGATATT CCTTGACATC CAAATTTAAA ACATCAAAAT AATTTTCCAT | 7320 |
| TGTAACATCT CTTAGTTCAA TTGTCATAGT TTTGCTCCTT GTTAGAGGTT ATCATTGGCG | 7380 |
| CAAAATAATG TTTACGGCAA ACTGAGATAT AGGTTCGTT ACCACCAATC TGGATCTGTT | 7440 |
| CTCCATCGTA AACGGGCACT CCATCCTGTG TTCGCAACAC CATGGTCGCC TTTTCTTGC | 7500 |
| AATACTGACA GATGGTCTTG ATTTCTGCAA TCTTGCTGTC TAAAAGCAAG AGATATTGCG | 7560 |
| AACCTTCGAA CAATTCATTG CGAAAGTCAT TTTTCAAGCC AAAAGCCATG ACGGGTATGT | 7620 |
| CTAACTCGTC CACAACACGA GCTAGGTCGT AAACATGGTG GCGTTTGAGA AACTGGGCTT | 7680 |
| CATCGACCAA AACACAGTAA GGTTTTCTG GTAGGTCTCG GATATAGCCA AAGATATCCG | 7740 |
| TTGTTCTCTC AATCGCAAGG GCAGGGCGTT TCATGCCAAT TCGACTCGAC ACATAGCCAA | 7800 |
| CGCCGTCACG CGTATCCAGA GCCGAGGTCA TAATCACAAC ACCTTTTCCT TGCTCCTCGT | 7860 |

713

| | |
|--|------|
| AGTTATAGGC CACTTTGAGA ATCTCAATCG TTTTACCAGA GTTCATGGTC CCATAACGAT | 7920 |
| AGTACAACTG TGCCATGTTT CTTGCTTCAC GTCCATTTCT AAATTTTTCG TACATTTCTAG | 7980 |
| TATATCATAA TTTTCTTAAG CTTTAAACGG CAAAATGTGG TAAAATAGAA GAAATCAAAA | 8040 |
| ACTAGTGGAG GAAGCTATTA TGCCATTTGT ACGCATCGAT TTATTTGAAG GACGCACGCT | 8100 |
| CGAGCAAAAG AAAGCTCTTG CTAAGGAAGT AACGGAAGCA GTTGTCCGCA AACTGGAGC | 8160 |
| CCCTCAATCT GCTGTCCATG TCATCATCAA CGACATGCCA GAAGGAACTT ACTTCCCACA | 8220 |
| AGGGGAAATG CGTACTAAAT AAGCTAGCTT AAGCAGAATT GCTTAGGCTT TTTCAATCTC | 8280 |
| CAAGTAGCAT TCATTGAAGA AATATCCTAA ATTTGTTACA ATTTGAAAAG AACTTGGAG | 8340 |
| AATTTCCAAG AAAAGAGCTA TTAATTAAAG GAAACATTAT GATTACACGT GAATTTGATA | 8400 |
| CCATCGCTGC TATCTCTACT CCACTAGGTG AAGGGGCTAT TGGTATTGTC CGCCTGAGCG | 8460 |
| GAACAGACAG TTTTGCTATT GCGCAAAAGA TTTTTAAAG AAAAGACTTG AACAAGGTTG | 8520 |
| CCAGCCACAC TCTCAACTAC GGTCACTTA TTGATCCTCT GACTGGTAAA GTCATGGACG | 8580 |
| AGGTTATGGT TGGGGCTATG AAGTCTCCAA AGACCTTCAC TCGTGAGGAT ATTATCGAGA | 8640 |
| TTAACACCCA CGGTGGGATT GCGGTGACCA ATGAAATTCT CCAGCTAGCT ATTCGTGAAG | 8700 |
| GGGCTCGGTT GGCAGAACCT GGTGAATTTA CCAAACGTGC TTTTTTAAAC GGTGCGGTAG | 8760 |
| ACTTGACACA GGCAGAGGCT GTGATGGATA TCATCCGTGC CAAGACTGAC AAGGCCATGA | 8820 |
| ACATTGCGGT CAAACAATTA GACGGCTCCC TTTCTGACCT CATTACAAT ACCCGTCAAG | 8880 |
| AAATCCTCAA TCACTTGCC CAAGTTGAGG TCAATATCGA CTATCCTGAG TATGACGATG | 8940 |
| TTGAGGAAGC CACTACTGCT GTTGTCGAG AGAAGACAAT GGAGTTTGAG CAATTACTAA | 9000 |
| CCAACTCCT TAGGACAGCA CGTCGTGGTA AAATCCTTCG TGAAGGAATT TCAACGGCTA | 9060 |
| TCATTGGACG TCCCAACGTT GGGAAATCAA GCCTTCTCAA CAACCTCTTG CGTGAGGACA | 9120 |
| AGGCTATCGT AACAGATATC GCTGGGACAA CACGAGATGT CATCGAAGAG TACGTCAACA | 9180 |
| TCAATGGTGT ACCTCTCAAA TTGATTGATA CAGCCGGTAT TCGTGAAACG GATGATATCG | 9240 |
| TTGAACAAAT TGGAGTTGAG CGTTCGAAAA AAGCTCTTAA GGAAGCTGAC CTAGTTCTGC | 9300 |
| TAGTACTAAA CGCTAGTGAA CCACTAACCG CCCAAGATCG CCAACTCCTA GAAATCAGTC | 9360 |
| AGGAGACTAA TCGCATTATT CTTCTTAACA AAATGACCT GCCTGAAACG ATTGAAACTT | 9420 |
| CGGAACTACC TGAAGATGTC ATCCGCATTT CAGTTCTTAA AAATCAAAAC ATCGATAAAA | 9480 |
| TCGAAGAGAG AATCAACAAC CTCTTCTTTG AAAATGCTGG TTTGGTTGAG CAAGATGCTA | 9540 |
| CCTACTTGTC AAACGCCCGT CACATTTCTT TGATTGAGAA GGCCGTTGAA AGCCTACAAG | 9600 |

714

| | |
|---|-------|
| CTGTTAACCA AGGTCTTGAA CTAGGGATGC CAGTTGACTT GCTTCAAGTT GACTTGACCC | 9660 |
| GTACTTGGGA AATTCTAGGA GAAATCACTG GAGATGCTGC TCCAGATGAA CTCATCACCC | 9720 |
| AACTCTTTAG CCAATTCTGT TTAGGAAAAT AAGAAAAATC CATGATCCTT CATTCCGTCA | 9780 |
| TGGATTTTAG GTTCTATAAT ATTTGTAGTG GGTAAATCCA CTATAGATAT TATGGAGCCT | 9840 |
| ATTTTATGTG AGAAAAAAG TCCCATATGA CCTATAATGA AAAGCGACAA AACAACTCAT | 9900 |
| TAGAAAGAAT CATATGGAAC AATTACATTT TATCACAAA TTAGTAGACA TTAAAGACCC | 9960 |
| TAATATCCAG ATTTTAGACA TCATCAATAA GGATACACAC AAGGAAATCA TCGCCAAACT | 10020 |
| GGACTACGAC GCCCATCTT GCCCTGAGTG CGGAAACCAA TTGAAGAAAT ATGACTTTCA | 10080 |
| AAAAACCTTC TAAAATTCCT TATCTTGAAA CGACTGGTAT GCCCACTAGA ATTCTCCTTA | 10140 |
| GAAAGCGTCG ATTCAAGTGC TATCACTGTT CAAAATGAT GGTCGCTGAA ACTTCTATCG | 10200 |
| TCAAGAAGAA TCACCAAATC CCTCGTATCA TCAACCAAAA GATTGCTCAA AAGTTAATTG | 10260 |
| AAAAGATTTC TATGACTGAT ATTGCCCATC AGCTTTCCAT CTCAACTTCA ACTGTTATTC | 10320 |
| GTAAGCTCAA TGACTTTCAC TTAAACATG ATTTTCTTG TCTTCCTGAG ATTATGCTT | 10380 |
| GGGATGAGTA TGCTTTTACA AAAGGGAAGA T | 10411 |

(2) INFORMATION FOR SEQ ID NO: 90:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2393 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 90:

| | |
|--|-----|
| GTTTTGGGTT CTGGAAATTA TCAGATGGTT GGAAAAGCCG TCCACATCAA GATAGTGTTC | 60 |
| GGAGATTTAA GTTTAAATG AAGAACTAA CACAGAGGAA ATGGAGTATA GACCTAACAA | 120 |
| GACGTATTGA GCAACTGAAT TTGTCTATTG GAGGATGGAT AAACCTATTGC TCATTGGGAA | 180 |
| ATATGAAAAG TATAGTCGCC AGCATAGATG AGCGCTTGGC TACTCGCCTA CGAGTGATTA | 240 |
| TCTGGAAGCA ATGGAAGAAG AAATCGAGAC GATTATGGGG ATTGCTTAAG TTAGGAGTTC | 300 |
| CTAAATGGAT AGCAGATAAG GTATCTGGCT GGGGCGACCA TTATCAATTA GTAGCTCAGA | 360 |
| AGTCGGTACT TAAACGTGCT ATATCAAAAC CAGTCCTGGA AAAACGTGGA CTGGTTTCGT | 420 |
| GTTTGGATTA TTACCTTGAA CGACATGCGT TAAAAGTTAG TTGAACCGCC GTATGCCAAA | 480 |
| CGGCACGTAC GGTGGTGTGA GAGGGGCTAG AGATTATCCC CTAATCGATT AACTCCCCTG | 540 |
| AAATTTATTT TAATTATGCA AATTTACGT ATTTTGTATG CTGAGACGAC GATCCTGGGA | 600 |

715

| | |
|--|------|
| ACTTTTCAGA TATTTTTTTG ACTATCTAAA TCTATCATTA GAAAAGCTTA GAGCGCCAAA | 560 |
| GGATTGAGC GTTTTTCTGA TTTTAAAGAC TTTTCCAGT CTCTTTTTCG ATTGAAGATG | 720 |
| TAATTATTCT ACTAACTAAC TAACTCTTA GTACTAGCCA ACAACGATAA TCATAATTCC | 780 |
| TCCTAAAATT AGGAATAATA AAGGCAATAG TTTTGTITT TTCATGTAAA AAACCTCACT | 840 |
| TTTGTTTTCT GCTATTTTAT GCTAAAATAT TAAAAATCAA ATTTAATTCC AAAGTTTGTA | 900 |
| ACTAAAGGGG GAGCGCTACA TGTCTAATTC ATTTGTCAAG TTGTTAGTCT CTCAATTATT | 960 |
| TGCAAATTTA GCAGATATTT TCTTTAGAGT AACAATCATT GCTAACATAT ACATTATTTT | 1020 |
| AAAATCAGTA ATTGCCACAT CACTAGTTCC TATCTTAATA GGAATATCCT CTTTGTGTGC | 1080 |
| GAGTCTTTTA GTTCCGTTGG TTAATAAAG GTTAGCGCTA AATAGGGTTT TATCTTTATC | 1140 |
| TCAATTTGGA AAGACTATAT TATTGGCGAT ACTGGTAGGA ATGTTTACCG TAATGCAATC | 1200 |
| CGTAGCGCCT TTGGTGACCT ATCTATTTGT TGTGCAATT TCCATACTAG ATGGTTTGTG | 1260 |
| AGCACCCGTT TCCTATGCTA TTGTGCCACG CTATGCGACC GATTTGGGTA AGGCTAATTC | 1320 |
| AGCCTTATCA ATGACTGGTG AAGCTGTCA ATTGATAGGT TGGGGATTAG GTGGACTCTT | 1380 |
| GTTTGCAACA ATTGGTCTGT TACCTACCAC GTGTATCAAT TTAGTCTTGT ATATCATTTT | 1440 |
| TAGCTTTCTG ATGTTATTTT TTCCTAACGC TGAAGTGGAG GTGTTAGAGT CAGAACTAA | 1500 |
| TCTTGAAAT TTGCTCAAAG GTTGAAGTT AGTTGCTAGA AATCCTAGAT TAAGACTTTT | 1560 |
| TGTATCAGCA AATTTATTGG AAATTTTTC AAATACGATT TGGGTTTCTT CCATTATACT | 1620 |
| TGTTTTTGTA ACGGAGTTAT TAAATAAAAC GGAAAGTTAC TGGGGATATT CTAATACAGC | 1680 |
| ATACTCTATT GGTATTATAA TTAGTGGCTT AATTGCTTTT AGGCTATCTG AAAAGTTCCT | 1740 |
| TGCTGCTAAA TGGGAACCCC AATTATTAC CCCAAATCTA AAAACCATCC AGAATCCTTG | 1800 |
| CCTTAGCTTA GATCCTGGAT GGTTCCTTTT TTCACCCAAT GGGTGTTTTT TACTAGACAA | 1860 |
| AAAAGAGTTT CCCCTTTATG GTATAAGTGT AGAAAAAAC ACAAAAAGAA AGGAAACTCA | 1920 |
| CATGAACAGT TTACCAAATC ATCACTTCCA AAACAAGTCT TTTTACCAAC TATCTTTCTG | 1980 |
| TGGAGGTCAT TTAACCCAGT ATGGTGGTCT TATCTTTTTC CAGGAACCTT TTTCCAGTT | 2040 |
| GAAACTAAAA GAGCGGATTT CTAAGTATTT AGTAACGAAT GACCAACGCC GCTACTGTCTG | 2100 |
| TTATTCGGAT TCAGATATCC TTGTCCAGTT CCTCTTTCAT CTGTTAACAG GTTATGGAAC | 2160 |
| GGAATATGCT TGTAAGAAT TGTGAGCTGA TGCCTACTTT CCAAAATTGT TGGGAAGGAG | 2220 |
| GCAGCTTGCT TCACAGCCAA CCTTATCCCG TTTTCTTTCC AGAACTGACG AGGAAACAGT | 2280 |
| CCATAGTTTG CGATGCCTCA ACCTGAATT GGTGGAATTC TTTTACAGT TTCACAGCT | 2340 |

716

AAACCAACTC ATTCTAGATA ACGATCTAC CCATTTTACA ACTTATGGCA AGC 2393

(2) INFORMATION FOR SEQ ID NO: 91:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4762 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 91:

| | |
|---|------|
| TTGTATCTT TTTAGGTCTC TTTCAATCCA AACCCTTTAA ACTATACGTC ATTTTCGGTTC | 60 |
| CTGCAAGTCT TGTGGTAATT TTAGGTTTGA TTTTACTTTT CTTTTCACAA GAGCCTCTGC | 120 |
| ACGCTTCTTA TTTGATGGTC GTCTTCCCTG TTTTCTACT TTTATTGGTA ACCAATATTA | 180 |
| AGAGTCAACA GAGGGGGCGT AGTGCTAGAA GAAGCCGAAG AGAAACGCCA TTATGCCTAT | 240 |
| GGAGTCGTTT CTTCAAAGGA AATCTATATC TGCTAGTTT TGGGTTTGTC TATCTTTTGT | 300 |
| CTGTTCCTTT TTTGATGAAG TTTGTCCTTT ATCCAGTACC TTATCAAGAA CGTAATCGTC | 360 |
| TTGCTGATTT GGTAAAAGAG GAGACAAATA CGGAAGATGC TATCTCATGC ATGGGATGAT | 420 |
| ACTGCGACTC TTTATCGTAA GAGTGAGCGC TTGTCCCATC GCGGATTTTG TCCCCGTGTC | 480 |
| ACTATACAGC AACTGAGGAA AATCGTAATA AGTTACTTAA TGACTTGAAA GAAAAACAAC | 540 |
| CTAAGGTGAT TGTGGTAAAT GATAAGGTGG TAGTCTGGTC TGAAGTGGA ACACCTCTTA | 600 |
| AAGAAAATTA CCAACAAGTA AAGACTGATT ACTCAGAGTT TAAAGTCTAT AAAATTAAT | 660 |
| AACCAATCA ATATCTTGTG TATTTTAAA AATTTTAGGA TTTTAAACAC AAGATATTGA | 720 |
| TTTTTCCTTT TAGAGTGGTA TAATACTTTT TAGAAGAAG ATTTTAGAAA AGAGCATGCA | 780 |
| TATGATTGCA CTAGAAGAAA AAATTACAAT TTTGCCAACT CTCTTCGTCG AGAAACGAGA | 840 |
| TGGGAGACGT GTTGTATTG ATGTGGACAA GATTGACAAG GCTCTCCACA AGGCGGCTGA | 900 |
| CAAGGTTATG GATGTGACAC CCCTGGTTGA AAAATGCCTC AATGATCTGA CTGAGCGAAT | 960 |
| TATTACAGAA ATTCATAGTC GCTTCCACA GGAATTAAG ATTTACGAAA TTCAAAATAT | 1020 |
| CGTAGAACAT GAACTCCTTG AAGCCAAAGA ATATGCGCTG GCTGAGGAGT ATATTACTTA | 1080 |
| TCGGACACAG AGGGATTTTG AGCGCTCAAA AGCGACGGAT ATCAACTTTA GTATTCATAA | 1140 |
| ACTTCTCAAC AAAGACCAGA CAGTTGTCAA TGAAAACGCT AATAAAGACA GTGATGTCTT | 1200 |
| TAACACTCAG CGTGATTGTA CAGCAGGGAT TGTTGGGAAA TCAATCGGAC TGCAAATGCT | 1260 |
| TCCTAAGCAC GTAGCCAATG CCCACCAAAA GGGGATATC CACTATCACG ATTTGGACTA | 1320 |
| CAGTCCCTAT ACCCCTATGA CCAACTGCTG TTTGATTGAT TTTAAGGGTA TGTGGAAAA | 1380 |

717

| | |
|---|------|
| TGGTTTAAAG ATTGGAAATG CAGAGGTAGA GAGTCCCAAG TCTATCCAGA CTGCGACAGC | 1440 |
| ACAGATTTCT CAAATCATTG CCAACGTTGC TTCTAGCCAG TACGGTGGCT GTTCAGCTGA | 1500 |
| CCGTATCGAT GAAATTTTGG CGCCTTATGC AGAGAAGAAT TATCAAAAAC ATCTCAAAGA | 1560 |
| TGCAGAAGAG TGGGTATTGC CTGAAAAACA GGAAGATTAC GCTTGGAAGA AAGCGCAAAA | 1620 |
| GGACATCTAC GATGCCATGC AATCTCTTGA GTATGAAATC AATACTCTCT TCACTTCAAA | 1680 |
| TGGACAAACA CCTTTTACTT CGTTAGGTTT TGGTCTGGA ACCAGTCGTT TTGAACGAGA | 1740 |
| AATCAAAAA GCTATTTTAA ACATTTCGCAT CAAGGCTCTT GGTTCAGAAC ACCGTACGGC | 1800 |
| TATCTTTTCT AAACCTATCT TTACGCTTAA AAGAGGCCTC AACTTAGAGG AAGGAATCC | 1860 |
| CAACTATGAC ATCAAGCAGT TGGCTCTAGA GTGTGCAACC AAGCGGATGT ATCCAGACGT | 1920 |
| CTTGTCTTAT GATAAGATTG TTGATTTGAC AGGTTCTTTC AAGGTGCCTA TGGGCTGCCG | 1980 |
| TTCTTTCTCT CAAGGGTGGA AGGATGAAAA TGGTGTAGAA GTCAATTCAG GTCGCATGAA | 2040 |
| TCTGGGTGTT GTGACGGTTA ATCTGCCTCG TATTGCTCTT GAGTCTGAAG GTGATATGAA | 2100 |
| TAAGTCTCG GAAATCTCA ACGAGCGAAT GAATATCGCA GAAGATGCTC TTGTTTACCG | 2160 |
| TGTCGAACGC ACTAAAGAGG CGACACCAGC GAATGCTCCT ATTCTTTATC AGTACGGTGC | 2220 |
| TTTTGGCCAT CGTCTAGGTA AAGAAGAAAG TGTTGACCAG CTCTTTAAGA ATCGTCGTGC | 2280 |
| GACCGTTTCG CTGGGCTATA TCGGCTTGTA TGAAGTAGCG ACAGTTTCTT TTGGTAACAG | 2340 |
| CTGGGAAAGT AATCCAGATG CTAAGGAATT CACGCTAGAC ATCATTCACG ATATGAAACG | 2400 |
| CCGTGTAGAA GAGTGGTCAG ACCAATATGG CTACCATTTT TCTATCTACT CAACACCATC | 2460 |
| CGAAAGTCTG ACAGACCGTT TCTGCCGACT AGATATAGAC AAGTTTGGCT CTATTCCTGA | 2520 |
| TATCACAGAC AAGGAATACT ACACCAACTC TTTCCACTAC GATGTTGCTA AAAATCCAAC | 2580 |
| ACCGTTTGAA AAATTGGACT TTGAGAAAGT CTATCCGGAA GCAGGTGCGT CAGGTGGTTT | 2640 |
| CATCCATTAT TGTGAGTATC CAGTCCTTCA GCAAAATCCA AAGGCCTTGG AAGCTGTCTG | 2700 |
| GGATTATGCT TATGACCGTG TAGGCTATCT AGGCACCAAT ACTCCGATTG ACCGTTGCTA | 2760 |
| CAAGTGTGAC TTTGAAGGGG ATTTTGAACC AACTGAGAGA GGGTTTGCTT GTCCAAACTG | 2820 |
| TGGCAATAGC GACCCATAAA CAGTAGATGT GGTGAAACGA ACTTGTGGCT ACCTAGGTAA | 2880 |
| TCCTCAAGCA AGACCGATGG TCAACGGGCG TCACAAGGAA ATCGCTGCGC GTGTCAAACA | 2940 |
| TATGAATGGT TCAACGATTA AAATAGCTGG GCATCAAGTA ACAAATTAGA AAGAAATGAA | 3000 |
| ATGGGAAAA ATCAACTAGA CGATAAGGGG CGCGCACAAG TGACCCGTTA TCACGAGAAA | 3060 |
| CACTCTAAAG GTGAGCTGG TAAGAAAGAA CGCTTGCTTA GCTTCAGAGA ACAATTTTAA | 3120 |

718

| | |
|---|------|
| AACAAGAACA AGAAAAAATA AAAGTGAGAG CCAGCTCTCG CTTTCTCAT AGTGGGAGGT | 3180 |
| AAGGATGGAA TTACGCAGAC CAAGATTAGC GGATAAGAAA GCTGTTTTAG ATATGATGAC | 3240 |
| AGAGTTTGAA AAATTCAGT CGCCTCACGA CGGCGGTTTC TGGGATACAG AGAACTTTGT | 3300 |
| GTATGAAGAC TGGTTAGAAA GCAATCAGGA ACAGGAAATG GGGATTAATC TGCCTGAAGG | 3360 |
| ATGGGTTTCT GCAATTCAGT TAGTGGCTTT TTCTGAGAAA GGTCAAGCAG TTGGATTTCT | 3420 |
| TAATCTCCGG TTGCGCCTCA GTAACTTTCT ACTAGAAGAA GGTGGCCACA TTGGCTACTC | 3480 |
| CATTCTGCCA TCTGAAAGAG GCAAGGGTTA TGCAAAAGAG ACTCTCCGTC AGGGCTTGCA | 3540 |
| AGTTGCTAAG GAAAAGAACA TCAAGAAAGC TCTGGTGACC TGTAGTGTGA ATAATCCTGC | 3600 |
| TAGCAGAGCA GTCATTCTAG CAAATGGTGG AATATTTGAG GATGCTCGCA ATGGAGTCGA | 3660 |
| GCGTTATTGG ATAGAGGTAG CGAATGAATA ATCCAAAACC ACAAGAATGG AAAAGCGAGG | 3720 |
| AACTTAGTCA AGGTCGTATC ATTGACTACA AGGCCTTTAA CTTTGTGGAC GGCGAAGGCG | 3780 |
| TGCGCAACTC TCTCTATGTA TCAGGCTGTA TGTTTCACTG CGAGGGATGT TATAATGTTG | 3840 |
| CGACTTGGTC TTTTAATGCT GGCATTCCCT ATACAGCAGA ATTAGAAGAG CAGATTATGG | 3900 |
| CAGACCTTGC CCAACCCTAT GTTCAAGGCT TGACTTTGCT GGGAGGGGAG CCTTTTCTCA | 3960 |
| ATACTGGGAT TCTCTTGCCA CTTGTTAAGC GGATTCGGAA GGAATTGCCA GACAAGGACA | 4020 |
| TCTGGTCCGT GACCGGCTAC ACTTGGAAG AAATGATGTT GGAAACTCCA GATAAACTGG | 4080 |
| AATCTTTGTC ACTGATTGAC ATTCTTGTG ATGGAAGATA TGATCGAACT AAGAGAAATC | 4140 |
| TTATGCTCCA GTTTCGAGGT TCATCTAACC AACGAATTAT CGATGTGCAA AAATCGCTCA | 4200 |
| AAAGTGGGCA AGTAGTGATT TGGGACAAGC TCAATGACGG AAAAGAAAGC TATGAACAGG | 4260 |
| TGAAGAGAGA ATGAAGAAAA AGGACTTAGT AGACCAACTA GTCTCAGAGA TCGAGACGGG | 4320 |
| GAAAGTCAGG ACACTGGGAA TATACGGTCA TGGAGCTTCA GGTAAATCAA CCTTTGCACA | 4380 |
| GGAATTGTAC CAAGCTTTAG ATTCTACTAC AGTAAATTTG CTAGAGACAG ATCCTTATAT | 4440 |
| CACCTCAGGA CGCCATCTGG TAGTACCCAA GGACGCGCCG AATCAAAAGG TGACAGCCAG | 4500 |
| TCTGCCAGTG GCGCATGAAC TGGAGAGTTT GCAGAGAGAT ATCCTTGCTT GCAGGCGGGT | 4560 |
| ATGGATGTCT TGACAATTGA AGAACCTTGG AAGGCTAGTG AGGTCTTGTC TGGAGCCAAA | 4620 |
| CCAATTTTGA TTGTCGAAGG GATGTCTGTT GGCTTTCTAC CCAAGGAACT CTTTGAAAAA | 4680 |
| ACCATCTGTT TCTACACGGA TGAGGAGACC GAATTAAAGC GACGCCTTGC TAGAGATACC | 4740 |
| ACTGTGAGAA ATCGCGATGC GG | 4762 |

(2) INFORMATION FOR SEQ ID NO: 92:

(i) SEQUENCE CHARACTERISTICS:

719

(A) LENGTH: 3832 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 92:

| | |
|---|------|
| GATGCAGGTT TCGACCCACA TATTCCAGAA AATTACTTTA AAGATGATGA TGTTAATCAG | 60 |
| GTACCTTGTC TTGTTGGTC TTCATCTGCA GCCCTCTTTT TCAGTAATG GGTAGACCAT | 120 |
| GCGGTCTATC AGGAGACGCC TTTTGATTGG AGAAAGATAG AAGATGATGC ATCTGCATAT | 180 |
| GGGTATTTAT AAGAGGAATT ATGACATATT TAGACGCTTT TAAATCAGGT ACCTTGCTTT | 240 |
| TACCGAGTGC CCTGCTCTTG CATTTTAAGG AACTCTTTCC TTCTAGCGAC GATTTTCTGG | 300 |
| TTTGGCAATT TTTCTATTTG CAAAATACGA CAGGCTTAGA AGAAATGTCG CCAAGCCAGA | 360 |
| TTGCTGAAAG GATGGCAAG GAAATTTCGG ATGTCAACCA GTCCATTTCT AATCTGACGG | 420 |
| AAAGGGGACT GCTCCAGTAT CGTACTATCG AATTAAATGG CGAAATTGAA TTGCTCTTTG | 480 |
| ATGCTAGTTT GGCCTTGGA CGTTTGATG ACCTGTTTGG AGCAGTTCAT TCAAGTTCAG | 540 |
| ACCAGCTAAC ACCTCAAAAC CAGCTCAAGG ATTTGGTGGA AACCTTCCAG CAGGAGTTGG | 600 |
| GACGATTGTT GACGCCTTTT GAGATTGAGG ATTTGACCAA GAACTAAAG GAAGATGGAA | 660 |
| CCAGTGCTGA CTTGATTAAG GAGGCTCTTC GTGAAGCTGT TTGAATGGA AAACCAAAT | 720 |
| GGAAGTACAT TCAGGCGATT TTGAGAACT GCGCCATGA AGGAATCAAG AGTGTGGCTC | 780 |
| AAATTGAGGC CAAGAGAGCA GAAAGAGAAG CAAGCAATCC TCAGTTGACA CAGGTATCTG | 840 |
| CAGATTTTAT AAATGCCATG GATCTCTGGA AGGATTAATC CATGCAAGTA GGCTTGAAAT | 900 |
| CCGAGTAAGA TTTGCAAGCT GTGTATAATT GTGATAGAAT AAATAGAAAA TAAATTGAAA | 960 |
| AAAGAGGTAT GTGAAATGTC ACGTAAACCA TTTATCGCTG GTAAGTGGA AATGAACAAA | 1020 |
| AATCCAGAAG AAGCTAAAGC ATTCGTTGAA GCAGTTGCAT CAAACTTCC TTCATCAGAT | 1080 |
| CTTGTTGAAG CAGGTATCGC TGCTCCAGCT CTTGATTTGA CAACTGTTCT TGCTGTTGCA | 1140 |
| AAAGGCTCAA ACCTTAAAGT TGCTGCTCAA AACTGCTACT TTGAAAATGC AGGTGCTTTC | 1200 |
| ACTGGTGAAA CTAGCCCACA AGTTTGTGAAA GAAATCGGTA CTGACTACGT TGTATCGGT | 1260 |
| CACTCAGAAC GCCGTGACTA CTTCATGAA ACTGATGAAG ATATCAACAA AAAAGCAAAA | 1320 |
| GCAATCTTTG CGAACGGTAT GCTTCCAATC ATCTGTTGTG GTGAATCACT TGAACTTAC | 1380 |
| GAAGCTGGTA AAGCTGCTGA ATTCGTAGGT GCTCAAGTAT CTGCTGCATT GGCTGGATTG | 1440 |
| ACTGCTGAAC AAGTTGCTGC CTCAGTTATC GCTTATGAGC CAATCTGGGC TATCGGTACT | 1500 |

720

| | |
|---|------|
| GGTAAATCAG CTTACACAAGA CGATGCACAA AAAATGTGTA AAGTTGTTTCG TGACGTTGTA | 1560 |
| GCTGCTGACT TTGGTCAAGA AGTCGCAGAC AAAGTTTCGTG TTCAATACGG TGGTTCTGTT | 1620 |
| AAACCTGAAA ATGTTGCTTC ATACATGGCT TGCCCAGACG TTGACGGTGC CCTTGTAGGT | 1680 |
| GGTGCGTCAC TTGAAGCTGA AAGCTTCTTG GCTTTGCTTG ACTTTGTAAA ATAATCAGTA | 1740 |
| AGTAGCAAAA GCTAGGTGGA ACAGCATTCA GATGCTCTGTT ACATTTTTTA TAGGAGAGAA | 1800 |
| AGATTGAAAA CAAAAATTGG ATTAGCAAGT ATCTGTTTAC TAGGCTTGGC AACTAGTCAT | 1860 |
| GTCGCTGCAA ATGAAACTGA AGTAGCAAAA ACTTCGCAGG ATACAACGAC AGCTTCAAGT | 1920 |
| AGTTCAGAGC AAAATCAGTC TTCTAATAAA ACGCAAACGA GCGCAGAAGT ACAGACTAAT | 1980 |
| GCTGCTGCCC ACTGGGATGG GGATTATTAT GTAAAGGATG ATGGTTCTAA AGCTCAAAGT | 2040 |
| GAATGGATT TTGACAACTA CTATAAGGCT TGGTTTTATA TTAATTCAGA TGGTCGTTAC | 2100 |
| TCGCAGAATG AATGGCATGG AAATTACTAC CTGAAATCAG GTGGATATAT GGCCCAAAAC | 2160 |
| GAGTGGATCT ATGACAGTAA TTACAAGAGT TGGTTTTATC TCAAGTCAGA TGGGGCTTAT | 2220 |
| GCTCATCAAG AATGGCAATT GATTGGAAT AAGTGGTACT ACTTCAAGAA GTGGGGTTAC | 2280 |
| ATGGCTAAAA GCCAATGGCA AGGAAGTTAT TTCTTGAATG GTCAAGGAGC TATGATGCAA | 2340 |
| AATGAATGGC TCTATGATCC AGCCTATTCT GCTTATTTTT ATCTAAAAATC CGATGGAAT | 2400 |
| TATGCTAACC AAGAGTGGCA AAAAGTGGGC GGCAAATGGT ACTATTTCAA GAAGTGGGGC | 2460 |
| TATATGGCTC GGAATGAGTG GCAAGGCAAC TACTATTTGA CTGGAAGTGG TGCCATGGCG | 2520 |
| ACTGACGAAG TGATTATGGA TGGTACTCGC TATATCTTTG CGGCCTCTGG TGAGCTCAAA | 2580 |
| GAAAAAAAAG ATTTGAATGT CGGCTGGGTT CACAGAGATG GTAAGCGCTA TTTCTTTAAT | 2640 |
| AATAGAGAAG AACAAGTGGG AACCGAACAT GCTAAGAAAG TCATTGATAT TAGTGAGCAC | 2700 |
| AATGGTCGTA TCAATGATTG GAAAAAGGTT ATTGATGAGA ACGAAGTGGG TGGTGTCAAT | 2760 |
| GTTCTGCTAG GTTATAGCGG TAAAGAAGAC AAGGAATTGG CGCATAACAT TAAGGAGTTA | 2820 |
| AACCGTCTGG GAATTCCTTA TGGTGTCTAT CTCTATACCT ATGCTGAAAA TGAGACCGAT | 2880 |
| GCTGAGAGTG ACGCTAAACA GACCATTGAA CTTATAAAGA AATACAATAT GAACCTGTCT | 2940 |
| TACCCTATCT ATTATGATGT TGAGAATTGG GAATATGTAA ATAAGAGCAA GAGAGCTCCA | 3000 |
| AGTGATACAG GCACTTGGGT TAAATCATC AACAAAGTACA TGGACACGAT GAAGCAGGCG | 3060 |
| GGTTATCAAA ATGTGTATGT CTATAGCTAT CGTAGTTTAT TACAGACGCG TTTAAACAC | 3120 |
| CCAGATATTT TAAAACATGT AAATGGGTA GCGGCCTATA CGAATGCTTT AGAATGGGAA | 3180 |
| AACCCCTATT ATTCAGGAAA AAAAGGTTGG CAATATACCT CTTCTGAATA CATGAAAGGA | 3240 |
| ATCCAAGGGC GCGTAGATGT CAGCGTTTGG TATTAAGCGA TGATTTGAAA GAGGGATGTG | 3300 |

721

| | |
|---|------|
| ATAGTAGCAC CCTCTTTTTC TTTGTTTTAT GATAGTTCAT CCTCGAGTAA ATTCAAGTTC | 3360 |
| TTGCTCGGAA ATGAAGCTTA TATAGTAGAT TGAATATAGA CAAATACCTT GTGATTGGTA | 3420 |
| AAACATTTTA GAAATTCATT TACCTTTCCT AATCGACTTG GTTTCATCTT ATTTCAATCT | 3480 |
| ATTATAGTAT TGGGGAATTT CTTCAAACCA CATCAGCTTG GTCAGTTCTA CCTGCGACCT | 3540 |
| CAAACTTGT GCTTTGGTCA AGCTGGGTTT AGTTTCCTAG TTTGCTGATG GATTTCCATT | 3600 |
| GACTATAAGC ATCCAACCCT CTTTTGTCT TCTAAAGAAT TCTTAAATTA TCAGTCTATT | 3660 |
| GCAACTTTTC TCATATAAGT TCTTTGTCTT GCTATTGGTT TTCCTTAGTA GTATACTAAG | 3720 |
| GTAGTAATCA TTAAGAAGTG GTTACAAAA ATAATGAATG AGGTAAAGAA AATGGTAGAA | 3780 |
| TTGAAAAAAG AAGCAGTAAA AGACGTAACA TCATTGACAA AAGCAGCGCC GG | 3832 |

(2) INFORMATION FOR SEQ ID NO: 93:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 10690 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 93:

| | |
|--|-----|
| TGAAAAATC CTCATGAACC TGGCGCCAAT AGACAAGTGT CTTGTTTCCC TCACCTTCCT | 60 |
| TATAGGCATG GTCAGCTGAC ACTCGATTGA AGGGTTTAAC AGAAACCTTT GTAATTTCTGA | 120 |
| CAATGCAGAC AGCCTGATTT TGAATATCTA AAATGACATC GAAGGTCCCT ACTTGGGGAA | 180 |
| GTGGTTCGTC TTCTAGCACA TAGAGGTCAT AGGCTGATGC TGTGCTGTC TTTTCTCCTT | 240 |
| TAAACACCAA ATCCGCTAAA AGGTCTGGTT CAACTCCAAA AGCCCAGGCA TCGATTTTCAT | 300 |
| CTCCGATCAA AGGATTGATT TGCTTGTATT TATTCCACAT TTCTTGCGGT ATCATGGGTG | 360 |
| CTCCTTTGTA ATTTTITACT TTCTTCTTTT ATGTGTTTAA GATGATCTGG ATGGTCAATC | 420 |
| TCTAAATCAA AAATCTCTGG AATAGAACTG TAGTGGATAA TGCACTTGAT ACCCAAATGA | 480 |
| TTCAATTTTT GTATGAAAGA AGTATTCAGA TAGCCTGCTA CAGCAAAATC AATCTTGTTC | 540 |
| TTTCTTGCTT TATCCTGCAT ATCTCTTAGC ATATCTAACA TTATTGGACT TTCCATATCA | 600 |
| TGCCATTGAC TGTTTCTCAT AGTCGCAAAA ACAAAGGAAG TCAAATCATT CATTCCAACT | 660 |
| ACAACTTTTG AAATGCCCGT TTCCAGTATA CTAGATAAGT CAAAATACGC TGACGGTAAT | 720 |
| TCAATCATCG TTCCGACTTT CCCAGTAAAA CCCTGCTGAC GCAATACTGT AATAGCTTGT | 780 |
| TTTAATTGGT CGGCATCATT GACAAAAGGA AAGATAACAG ATAGATTGGG GTTGGTTTGA | 840 |

722

| | |
|---|------|
| TAAACTTCTG TAACGACATG TGCTTCAGCC TGAAATTCAT CCAAACACGC CAGTAAACGC | 900 |
| CTAGTTCCTC TATAGCCAAA CAAGGGATGC CCTTCGTCAA AAAACTCTTT AGTCCCCACT | 960 |
| AAACAATTGG CTTCTGTATT CGTTAATTCA GTAAAACGAT ACCAAACTTC CTTACCTAAG | 1020 |
| TAAAAGGAGC AAATAGTATC AAGATAATCT TTCACAAATT CCTGACAACT TTGTAATAGT | 1080 |
| ATATTTTGAT TGAGCTCTCT CAATAAGTAT TCCCCACGAA TCATGCCGAC GTGGTGAAAT | 1140 |
| AGTTGAGGAT AAATTTTTC AAGAATTTT TCGCCACTAA GGGCAAGTTG ATTTCTCATC | 1200 |
| ATTCACCTTC CAATTCATGT AAGAAGTCTT GTCCAGTTCT GGAAATCCTA ATAATTCAGA | 1260 |
| CTTAACCTTC AAGACTAATG GCGATGCATT TTCTTCTGTA ATCTCTTGAA TATCCATCCA | 1320 |
| AATATATCCA AGTGAATCAT TCGCACCATC AGACACAGCT TCCGAAATCG TAACTTGAGG | 1380 |
| TGCACTCTCA TTCATTTCAA CATCATACAA GGCTATGACA TGGTGAACCA TAAAATTTT | 1440 |
| TAACTCTTCC CTGACGAAAA CATCGTAGAT TCGAGGATTA GAGTAGCTTC TAACAGTAAA | 1500 |
| TCCCGTCTCT TCCATAACTT CTCTAGTCAG CGTTTCCGTC AGTCCTTCAC CAAGTTGCTG | 1560 |
| ACTGCCCTCA GGTAGATCAT ACCGATGTTG ATAAGGGCT CTCGTTTTT CAATGCAAAG | 1620 |
| TAACTTTCCA TTTTCAAAGC AAACACAGTA GACCCCAAAG TGATTTTGA TTTCCATCCA | 1680 |
| ACTCCTCCTA CTTCAAAGAC CAGCCACCAT CTATTGTCAA GATTTGTCT TGCATGGCGC | 1740 |
| TCGCTTTTCC ACTTGCTAAA AAAAGACTAA GCTCTGTAT TTCTCTGGC TCAATCCAGC | 1800 |
| GCTTGATTGG GGTTCACATA GCCACCCAGT CAGCCAAACC ACCTGGTTCA AAATCCGCAG | 1860 |
| CGGTCATAGC TGTCTGACT GCTCCTGGAG CGATACCAA GACCTGAATC CCAGCTTCAG | 1920 |
| CATAGTCTAG AGCCAACCTGC TTGGTGAAGC CAGCCAAGGC ATGCTTGAT GAAGTATAGG | 1980 |
| CGTGACCACC TCCACCTGCT AGGCTAGAAG CAATGGAACA CATATTGATG ATGATTCCT | 2040 |
| TTTTATTTTC CAGCATTTGT GTCAAATAAT ACCGAGTCAA CTCTACTGGA ATAATGTAGT | 2100 |
| TGATTTCAA AATCTCTTGA ATGTCCTGCG CCGTTTGTTC CAACAGTGGT TTGTAATCAT | 2160 |
| CCAAAACCTC AGCAGTATTA CACAAAACAT CCACCTGAGG GCACCAGTCA AAAATAGGTT | 2220 |
| CCAAGTCCAA GGTCAAATCT CTCTGTAAAA AGCGAAAATC ACCCTCTAAG AGTGCTTTT | 2280 |
| CACCTTGGTC AACTCCATAA ACTTGATAGC CCTTCTCTAA AAAGAGGCGA GCTTGAGCCA | 2340 |
| ATCCGATCCC TGAACCTACT CCGTAATGA GTACACGTTT AGTCATGCAC TTCTACCCAA | 2400 |
| TCCGTTGCCA AAACATCACA AACTGTCGGG CTCCACATGG AAAAACCTTC TCCTTCGCCA | 2460 |
| GAAACGTTGA TTAGGAAATA AGGTGTCATT TCAAGTGCAA GCCCATTTT CTCGATGGTA | 2520 |
| TCAAAGAGTT GGACATAGTT TTCCGCACCT CCCCAACCAG TTCGTACATA TTTTCTCTTA | 2580 |
| GCCTTTAACC CAGGCAGGAT CTCTTCAAAT GTCATGTTTT TCTCCTTAA TTCTACATTC | 2640 |

723

| | |
|--|------|
| TTCATTTAAT TATAGCAAAA AACCGCTTTA TACGGCTTTT TGAATGTGAG TTATTCAAAC | 2700 |
| CTGCTACTAC TTACGGCAAA TTATTCCTTG CAGCAAGATA AATTTCATAC CATTCCTTTTC | 2760 |
| TTGTTAAGCT AAAGTTTGCC GCTCGGCTAA CTTCTCTCAA GTGCTTAGGA TTTGTTGTAC | 2820 |
| CTACGACTGC CTGCATTTT GCTGGATAAC GCAATATCCA AGAAATGGCA ATAGTTGAAG | 2880 |
| AGGTTACTCC ATATTTAATA GCTAAACGAT CAAGTACTTG ATTTAAAGCT TGAAATTTCT | 2940 |
| CATTTC AAC AAAATTCCTT TTAATAATACC CGAATTGTAA GACAGACCAT GCTTGAATGA | 3000 |
| CCACATCGTG TAATTGGCAA TATTCAAAAA TGCTGCCATC TCGCATAGCT GCTTGACTAT | 3060 |
| CTTCCATATT AACATGAAAA GCTGATTCAA ATCCTGGAGT AAAAGCCGCA CTC AATTGTA | 3120 |
| GCTGATTAA AGCTAACGGC TGCTTGACAT CTTTMTTAAG CAACTCCATC ATCATAGGAT | 3180 |
| TTTGATTAGA AACTCCAAAA TCTCGAAGTT TACCTTGTTT ATAAAGGAGA TTAAAGGCTT | 3240 |
| CTGCTACTTG GTCAGATTCC ATCAAAGCAT CTGGTCGATG AAGGAGCAAG CTATCTAGAT | 3300 |
| GATCAATCTT CAATCTTTGC AAAATACCGT CTAAGTATTT TATAATATAG TCCTTAGAAA | 3360 |
| AATCAAAATA GGTAATTTCT TCAATGCGAA TGCCACATTT GGAAGTGAATC CACATCTTTT | 3420 |
| CTCTTAAATC TGGACGATTT TTTAGGACAA GACCTAACAG TTCTTCACAA CGACCACGAC | 3480 |
| CATAAATATC AGCCAAGTCG AAGGCATTGA TTCCAACAGA AAGTGCTGTT TCTACAAGCT | 3540 |
| CTTCAACTTC TTTTACAGAT TTATCTTTTA TTCTCATCAT TCCGAGAACA ATTTCTGATA | 3600 |
| ATTCTTTGTC ATCTTGACCA AGAGTTATGT ATCTCATCAA ATTTTCTCTCC TTAAATTTCT | 3660 |
| AACATTCTTC CCTTCATTAT AACAAAAAAC CGCTTTGCAA CGACTTTTGT ACTATACTTC | 3720 |
| ACTCCATTTT ATCTTCTTAA ACCCAGCGAA CAAGACAAAG ATTCCAATAA AGAGGACAGC | 3780 |
| TAAAGGAATA ACTTTTGTA GGAACACATT TGAAATTTCC ATCCACTCAT AATAACGGAG | 3840 |
| CAGAGAACCC ACCACAAGAT GGGCAATAAT CATACTGACA AATGGACGAA AGACCCTTC | 3900 |
| TTTCCAATTC CAAATACCGA TAACTAGCGA AATCGTAAAG ACAGACAAAC TATCCAGGG | 3960 |
| AGCCGGAATA TAAAGGCTC CTTCTTGAT GAAGCTTGCC ATTCCTACAT ATCCTAAAAC | 4020 |
| AACTAGAAGA ACTATAGTCC CAACAACAAT GTAAGTGCCA ATTTTCATTT TAGGAGAATC | 4080 |
| TTGGACTAAA CTTCTTCGTA AAATTGTGGC CACAAGTCCA AATCCAATCA GAAAAATAAG | 4140 |
| AAGTTGCCCT AAAAATGTGA GCAAATTGAC TGTTAAGAGA GGACCTTTAG AAAATCACT | 4200 |
| TAGTAGTTGA TAATAACGTA ATACCGCCAG GACAAGAATT GGCGTCAAAA GGGACTCTTT | 4260 |
| GATAGAACTG CGAGGTGCTC CCTTGAGAAT CTCTTTCATT ATTTTCTTAG GATTCTTACC | 4320 |
| TAGATAATCC TCTGCACTCA TGCCATCTCG TTCTGCTTCT GAGAAATCTA GCATCATCAA | 4380 |

724

| | |
|---|------|
| ATAGATCTGC TCTCTGAGAT AGTCTTCATC ATAGAGAAAT CCAGCAAGAT TAAAACTTTC | 4440 |
| CCACAACTCC TCAAAATACT TTTGATTCTC CTCAGAAAAC TCATGTAGCA AAGCGCTTGT | 4500 |
| TTCTTCGTAA TACTTCATTT TCTTCATGGT TTAACCCCA TTCTTAATCC CTTCTACTTT | 4560 |
| TTGACTCAAA TCGTCCCATT GTTGCCAAA GACTGAGACA CGCTCTTCTC CTTCTTTCAT | 4620 |
| TAATGAAAA TACTTCCGAT CTGGACCATC TGGCGACGGG CGCATGTCGC CTCTTATCCA | 4680 |
| TTGATTTTTT TCTAACTTTT GCAACAAAGG ATAAATAGTT CCTGGAACGA TAGTATCAAA | 4740 |
| TCCAGCCTCT CGCAAAGTCT GAACCAACTC ATAACCATAC CGCTCTTTTT GACCAATCAT | 4800 |
| ATCCAAGACA CAACCTTCAA GAACACCTTT TAATAGCTGA GTTCTTTTCA TCACCTCTCC | 4860 |
| CTTCTAATCT ATTTTGTAAT ACCTACTAGT GACTTCACCT ATAGTATATC ACTTCTACAC | 4920 |
| TAGTTTGTA AGCATAATAG TTAATACTCT TCGAAAATCT CTTCAAACCA CGTCAGCGTC | 4980 |
| GCCCTACCGT ATGTATGGT ACTGACTTCG TCAGTTTCAT CTACAACCTC AAAACATGT | 5040 |
| TTGAGCTGA CTTGTCAGT TTCATCTACA ACCTCAAAAC AGTGTTTTGA GCTGACTTCG | 5100 |
| TCAGTTTCAT CTACAACCTC AAAACAGTGT TTTGAGCTGA CTTGTCAGT TTCATCTACA | 5160 |
| ACCTCAAAA CATGTTTGA GCTGACTTCG TCAGTTTCGT CTACAACCTC AAAACAGTGT | 5220 |
| TTTGAGCAAC CTGCGGCTAG CTTCTAGTT TGCTCTTGA TTTTCATTGA GTATAAATAA | 5280 |
| AAAAACAGAA CTAGCCTGAA CTAGTCCTGT CTACTTTTAC CCAATCACAC TTCCATTGG | 5340 |
| TACAGCTGGA TCAACTGTGA GAAGGGTTAA TTTGCCATCA TGTTCAGCTG AGAGAATCAT | 5400 |
| ACCTGGCTG ACATATTTTT TCATCATTTT ACGTGGTTTG AGGTAGCAA CGATTTGAAC | 5460 |
| TTTCTTGCG ACCAATCTTT GTTCATTGG ATAGTATTTT GCAATTCCTG AAAGAATCTG | 5520 |
| ACGATCTTCT CCATCACCAG CATCCAAGCG GAATTGAAGC AACTTATCTG AACCTTCTAC | 5580 |
| TTTAGACACT TCTTTGACTT CTGCGACACG GATTTCAACC TTGTCAAAGT CTTCAAACCT | 5640 |
| GATTTCATCC TTGTTTAGTT TGAGCTCAAC TTCGTCCGGA TTCCATTCTT TTTGACTGC | 5700 |
| TGGTTTATTG CCTTCCATTT GTTCCTTGAT ATAGGCGATT TCTTCTTCCA TATTTAGACG | 5760 |
| TGGAAAGATA GGTGTTCTT TGGCAACTAC AGTCACATCT GCTGGGAAGT CAGCCAACT | 5820 |
| CAAGTTTCA AGACTAGAAA CTTCTTCCAA ACCAAGTTGA GTCAAAACTG CAGGACTAGT | 5880 |
| TTCCATCATA AATGGTTCAA TCAAGTGAGC AACTACACGA ATGCTGGCTG CCAAGTGGCT | 5940 |
| CATGACACTT GCCAATTGGT CACGAAGAGC TTCATCCTTG GCCAAGACCC ATGGTGCGGT | 6000 |
| CTCATCGATG TATTTATTGG TACGAGAGAT CAGAGTCCAG ACTGCTTCAA GCGCACGTGG | 6060 |
| ATAGTCAACT GCTTCCATGT GTGTATGGAA GTCTGCGATT GATTGTWCTG CAACCTCAGC | 6120 |
| AAGAACATGA TCATATTCAG TCACACCTTC TACATAGGCA GGGATTTGTC CATCAAAGTA | 6180 |

725

| | |
|--|------|
| CTTATTAATC ATGGAAACCG TACGGTTAAG GAGGTTCCCA AGGTCATTAG CCAATTCATA | 6240 |
| GTTGATACGG CCGACATAGT CTTCAAGAGT AAAGGTTCCG TCTGAACCAA CTGGAAGGTT | 6300 |
| ACGCATGAGG TAGTAACGAA GTGGATCTAG TCCATAACGC TCTACCAACA TTTCAGGGTA | 6360 |
| AACGACATTC CCTTTTGA CTAGACATTTT TCCGTCTTTC ATGACAAAACC AACCATGGGC | 6420 |
| AATCAAACGA TCAGGTAATT TAACATCCAA CATCATAAGA AGGATTGGCC AGTAGATAGA | 6480 |
| GTGGAAGCGA AGGATATCTT TTCTTACCAT ATGGAAGACT GTTCCATTCC AGAACTTGTC | 6540 |
| AAAGTTACCA TGTTCGTCTT GAGCGTAGCC AAGAGCTGTC GCATAGTTAA GAAGGGCATC | 6600 |
| AATCCAAACG TAGACAACGT GTTTTGGATT TGATGGGACA GGCCTCCCC ATGTAAAGGT | 6660 |
| TGTACGAGAT ACCGCCAAAT CTTCCAAGCC TGGCTCGATG AAGTTGCGTA GCATTTTCATT | 6720 |
| AAGGCGACCA TCTGGCGTGA TAAATTCAGG ATGAGCTTTG AAAAATTCGA CCAAACGGTC | 6780 |
| TTGGTATTTC CTAAGGCGAA GGAAGTATGA TTCTTCAGAA ACCCATTCAA CCTCATGACC | 6840 |
| TGATGGAGCA ATACCACCAG TCACATTTCC AGCTTCATCA CGGAAAACCT CTGCCAGCTG | 6900 |
| GCTTTCTGTA AAGAATCTT CGTCTGATAC TGAATACCAA CCAGAGTATT CACCCAAGTA | 6960 |
| GATATCATCT TGAGCAAGTA AGCGTTCAAA GACTTGTGCG ACAACTTTTT CATGGTAGTC | 7020 |
| ATCAGTTGTA CGGATAAATT TATCGTATGA GATATCTAGT AATTGCCAGA GTTCTTTAAC | 7080 |
| TCCAACCGCC ATTCCATCAA CATAGGCTTG AGGTGTAATA CCAGCTTCTT CCGCTTTCTG | 7140 |
| CTGGATTTTC TGACCATGTT CATCAAGACC TGTGAGATAA AATACATCGT AGCCCATCAG | 7200 |
| GCGTTTGTA CGTGCTAGGA CATCACATGC GATAGTTGTG TAGGCAGAAC CGATATGAAG | 7260 |
| TTTCCCAGAT GGATAGTAAA TCGGCGTTGT AATATAAAAA TTTTTCAG ACATAATTTT | 7320 |
| TCCTTTCCAG GCAAATGAAA CCTGTTTTTC TAACACTTCA TTATATCACA TTTTAAATGA | 7380 |
| ATTTCAATAG GGAATCCAT ACAAAAACAA GATAGACGAG TGTCCATCTT GTTGATCTCA | 7440 |
| TTCATAACGA AGGGCTTCAA TTGGATCAAG TTTGATGCC TTGTTGGCTG GCAAGACTCC | 7500 |
| AAAAATCATA CCAACACTAG CCGAACTGC AAGACTAAAT AGGGCGACTG GGATTGATAC | 7560 |
| TCCAACCTCT ATACCTTCTA TTAAACCTTG CAGTAACAAA CCTGCTAAG CAGTTAAACC | 7620 |
| ACTTGCAATT GTCAAGCCAA TTAAGCCACC TAACAAGGTC AAAATCATGG ATTCAATCAA | 7680 |
| AAACTGAATT AAAATATTGG CACGTGTTGC ACCCAAAGCC TTACGAAGAC CAATCTCACG | 7740 |
| AGTGCGCTCT GTCACCGAAA CCAGCATGAT GTTCATGACA CCAGTTCCTC CAACAAAGAG | 7800 |
| AGAAATCCCT GCGATGGAAC TAATAATCGT CGTCATAAAA CTAAACGATT GTTGAATTTT | 7860 |
| TGCAAATACA ACGGACTCAT CTGCCACCTG GTATTCTCCC TGTGTAAAGC CTGCAAGCTC | 7920 |

726

| | |
|--|------|
| TGTCATTTTT CGTGCCAGTT CTGGACCCAG AGTTGGGGTT AAACCTGGTAT CATTCACTCG | 7980 |
| AAAGACAATA TTAGCTATTT CATCTACATT AAAATTCGCA GCAAGGGAGA TATTGGTAGT | 8040 |
| AATAGGCAAG CCACCAAACC CATATATTTT TGATCTTTTA GCCTCCGGAC TAGTATAAAC | 8100 |
| CCCAATGACC CGGTAACATA ATCCATTGAC TTCTACAACC TTGTTAATAG CCTCTTGAGG | 8160 |
| AGATTCAAAT AAACATAATGG ACAATTCCTC ATCTAGCAAA ATGACACTTG CAAACTCTTT | 8220 |
| GAAATCTTGC TCTCTCAGAC TACGACCTGC AATAATTTCA TTCTTAACAG CGTCCATGTA | 8280 |
| AGTTCTGTTT CCACCTGTCA AATTAGCATT CTCAACCTTT TTATCTTGAT AGGTCAAGAT | 8340 |
| GGCATTCTGT GAATTGGTTA CATAGTAAT ATCCACTCCC TTCAGTTTAG CTGCCTCTTG | 8400 |
| GACCCAGGAT TCTTGCGGTT TTGGCGGTT AACAGGAACT TCCTCTTCCT TTCCAGAAAC | 8460 |
| CGTAAAGCT GATTGTTTCT GAGTAAAGA CCCGTCTTTA CTTTTTTTAG GAGAGAAAAA | 8520 |
| GACGCTAATA TTTTCTGAG ATTTAGTCAT ATCTTTATTG ACTTGACGAG ATAGGGAATC | 8580 |
| ACCCAAAGCC ATAATCACAA CAACTGATGA AACACCGATA ATAATCCCA TCATAGTAAG | 8640 |
| CAAAGAACGC ATCTTGAGAG CCATGATAGA TGAAAAGGCA AATTCAGAT TCTGCATCTT | 8700 |
| AGTTTTCTCT CTTTCCTAAC TGAGCACTGT CAGACGAAAT GACCCCATCC CGAATGACAA | 8760 |
| TCTGACGTTT GGCATAGGCA GCAATCTCAG GCTCATGCGT TACCATGATA ATGGTTTTTC | 8820 |
| CTTCTTTATT CAAATCAACC AATAATGCA TAATTTGGTT ACCTGTTTTG GTATCCAAGG | 8880 |
| CTCCTGTCGG TTCATCCGCT AGGATAATAG AAGGATTGTT TACCAAGGCA CGCGCAATGG | 8940 |
| CTACACGTTG CTTTGTGACCA CCAGATAATT CTGAAGGTAA ATGGTGAATA CGTCTGTCA | 9000 |
| ATTCAACCTT GTCTAAATAT TCCTCAGCCA ACTTGCGACG TTTTGAAGAC GAAACTCTG | 9060 |
| CGTAAATCAA GGGCAATTCT ACATTTTGCA GAGCATGAG CTTCGATAGA AGAAAGAACT | 9120 |
| GCTGAAAGAC AAAACCGATT TGTGTTTAC GGACCTTAGC TAGTTGTTTT TCACCAAGCC | 9180 |
| CAGCCACTTC TTGACCTTCA AGATAATATT CTCCACTGGT TGGTGTATCC AACATGCCAA | 9240 |
| TCGTATTTCAT CAGAGTGGAC TTACCAGACC CAGATGGTCC CATGATGGCT ACAAATTCAC | 9300 |
| CCTCATTCAC TTCTAGATTG ATATTTTGA GAACCTGCAG TTCTTGGTCA CCATTACGGT | 9360 |
| AACTTCTGAA GATATTTTTT AGACTAATTA GTTGCTTCAT CAGCCTTCAC CTCTTTTCCT | 9420 |
| TCTTCCAAGG AAGATGTTGG ATTACTGATG ACCTTAGCAC CGTTCGTTAA ACCAGAAGTG | 9480 |
| ATTTCTTGAT TTTCTGCGTC AGCATTTCCC AATGAAACCT CAACTTTTTT AGCCTTTTGT | 9540 |
| TGTTTCATCCA CAATCCAGAC ATAATTTTGA CTATCATCCA TTAGTAGACT GCTAACAGGA | 9600 |
| ACAAGAATAG CCTTAGTTT GCTTTTAACC TCAATGTTGA CAGAAAAACC TTGTTTCAAA | 9660 |
| TCACCAACCT CGCCTGTCAC ATCAATAGTA TAAGGTATT TAGAACCTGT ATTATTCCCCG | 9720 |

727

| | |
|---|-------|
| GCTGCTGGAC TAGCTGCTTC ACCATTGTTT TTAGGATAGT CAGAAATATA GCTTAATTTC | 9780 |
| CCAGTCCATT TTTTATCAGG ATACACTTTA GAAGTAAAGC TTACTTCTTG ACCTACAGAA | 9840 |
| AGGTTGGCTA GATTGTACTC AGACAATTCT CCCTTGACTT GTAAATTTTC ATTGCTGACA | 9900 |
| ATATGAACCA TAACTTGACT CGCCCCTGTT GGAGATTTAG AAACATTGCT ATTGACTTCG | 9960 |
| ACCACAGTTC CCTCTAGGGT ACTGAGAACA GTTGTTCAT CCAATTGACT TTGAGCCTTG | 10020 |
| CTTAATTGCG CCGCAGCATC TGCACGCGCA TCACGGGCAT CACCCAATTG AGCGTCAATA | 10080 |
| GAAGCAACAG AATTTCAGC CACTGGAGTT GGGCTTTGCA CCGTTGCATC TTCTCCTCCT | 10140 |
| ACTGGCGCTG GTAAGTGTGG AGCCGGAGCT GAAGCGGCTT CATTTCTGTC TTGATTGAGT | 10200 |
| TCATTGATAT GACGATCTGC CCTAGCTACT GCTCGACTAG CTGAATCATA GGCCGCCTGC | 10260 |
| GCTTCTGAAC TACTGTACTT GACTAAAGCC TGCCCTTCGC TGACCTTATC GCCCACAGAA | 10320 |
| ACAAGGATTT CATCTAAATC ACCCTTACTA GCATCAAAAT AAACATATTG TTCATTTTTT | 10380 |
| GCTGTACTG TCCCTGACAA TAAACAGAG GAGGCCACGC TTCCTTCCTT GGCAACAACA | 10440 |
| AGATGAGTAG GCTCATCTTT TAGAGCAGTC TGAGAAGGTT GTCTAAAGAG TAAATCCCC | 10500 |
| CCAGCACCCA ATACAACTAC ACTCGCAGCA CCGATTGCTG CATACAGTTG CCACTTTTTA | 10560 |
| GCTTTACCAT TCTTTTCTT CATAATGAAA CTCCTTTCT TTTTACAAT ACTTGCTAT | 10620 |
| TATACCAAT TTCCCTCCAG CAAACAATAC AGTTCAGGAT TAAACAATCG TTCGGAATTT | 10680 |
| TGCTTTTCGG | 10690 |

(2) INFORMATION FOR SEQ ID NO: 94:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 8195 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94:

| | |
|---|-----|
| GAGAAAGCGC CCACGTTTCC CCGAAGGGAG AAAGGCGGAC AGGTATCCGG TAAGCGGCCA | 60 |
| GGGTCGGAAC AGGAGAGCGC AACGAGGGAG CTTCCCAGGG GGAAACGCCT GGTATCTTTA | 120 |
| TAGTCCTGTC GGGTTTCGCC ACCTCTGACT TGAGCGTCGA TTTTGTGAT GCTCGTCAGG | 180 |
| GGGCGGAGC CTATGGAAAA ACGCCAGCAA CGCGGCCTTT TTACGGTTCC TGGCCTTTTG | 240 |
| CTGGCCTTTT GCTCACATGT TCTTTCCTGC GTTATCCCCT GATTCTGTGG ATAACCGTAT | 300 |
| TACCGCCTTT GAGTGAGCTG ATACCGCTCG CCGCAGCCGA ACGACCGAGC GCAGCGAGTC | 360 |

728

| | |
|---|------|
| AGTGAGCGAG GAAGCGGAAG AGCGCCCAAT ACGCAAACCG CCTCTCCCG CGCGTTGGCC | 420 |
| GATTCATTAA TGCAGCTGGC ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA | 480 |
| CGCAATTAAT GTGAGTTAGC TCACTCATTA GGCACCCAG GCTTTACACT TTATGCTTCC | 540 |
| GGCTCGTATG TTGTGTGGAA TTGTGAGCGG ATAACAATTT CACACAGGAA ACAGCTATGA | 600 |
| CaTGATTACG AATTCGAGCT CGGTACCCGG AAAATCCAGA AAATGCTTGA AAAAAATCCT | 660 |
| AGAAGATGGT ATAATACTAA ATTGTAAGGG TTATCACATA TAACTCAAAA AAAGAAAGAA | 720 |
| CAAAAGGAGA GTCAAATAT GGCTTCTAAA GATTTCCACG TAGTGGCAGA AACAGGTATT | 780 |
| CACGCACGTC CAGCAACATT GTTGGTACAA ACTGCTAGCA AATTTGCTTC AGATATCACT | 840 |
| CTTGAGTACA AAGGTAAATC AGTTAACCTT AAATCAATTA TGGGTGTTAT GAGTCTTGGT | 900 |
| GTTGGCCAAG GTGCTGACGT AACTATCTCA GCTGAAGGTG CAGATGCAGA TGACGCTATC | 960 |
| GCTGCAATCT CAGAAACAAT GAAAAAGAA GGATTGGCAT AAGGGAAATG ACAGAAATGC | 1020 |
| TTAAAGGAAT CGCAGCATCT GACGGTGTG CAGTTGCAA AGCATATCTA CTCGTTCAGC | 1080 |
| CGGATTTGTC ATTTGAGACT ATTACAGTCG AAGATACAAA CGCAGAAGAA GCTCGCCTTG | 1140 |
| ATGCCGCTCT ACAGGCATCA CAAGACGAGC TTTCTGTTAT TCGCGAGAAA GCAGTAGGTA | 1200 |
| CGCTCGGTGA AGAAGCAGCT CAAGTTTTG ATGCTCACTT AATGGTTCTT GCTGACCCAG | 1260 |
| AAATGATCAG CCAAATCAAG GAACTATCC GTGCGAAGAA AGTGAATGCA GAAGCAGGTC | 1320 |
| TGAAAGAAGT TACAGATATG TTTATCACTA TCTTTGAAG CATGGAAGAC AACCCATACA | 1380 |
| TGCAAGAACG CGCAGcGGAT wTCCGCGACG TGCAAAAACG TGTATTGGCA AACCTTCTTG | 1440 |
| GTAaaaaaAT GCCAAACCCA GCTTCTATCA ATGAAGAAGT GATTGTGATT GCGCATGACT | 1500 |
| TGACTCCTTC AGATACAGCT CAATTGGACA AAACTTTGT AAAAGCTTTT GTAACCAACA | 1560 |
| TTGGTGGACG TACAAGCCAC TCAGCTATCA TGGCACGTAC ACTTGAAATT GCTGCTGTAT | 1620 |
| TAGGTACAAA TAACATCACT GAAATCGTTA AAGACGGTGA CATCCTTGCT GTTAACGGGA | 1680 |
| TCACTGGAGA AGTGATTATC AACCCAACAG ATGAACAAGC GGCAGAATTT AAAGCAGCTG | 1740 |
| GTGAAGCCTA TCGGAAACAA AAAGCTGAAT GGGCACTTTT GAAAGATGCT CAAACAGTGA | 1800 |
| CTGCTGACGG TAAACACTTC GAGTTGGCTG CTAATATCGG TACTCCAAAA GACGTTGAAG | 1860 |
| GTGTTAACAA CAACGGTGCA GAAGCTGTTG GACTTTACCG TACAGAGTTC TTGTACATGG | 1920 |
| ATTCTCAAGA CTCCCAACT GAAGATGAGC AGTATGAAGC ATACAAGGCT GTTCTTGAAG | 1980 |
| GAATGAACGG TAAACCTGTT GTCGTTCTGA CAATGGATAT CGGTGGAGAT AAGGAACTTC | 2040 |
| CTTACTTCGA TATGCCTCAC GAAATGAACC CATTCCTTGG ATTCCGTGCT CTTCGTATCT | 2100 |
| CTATCTCTGA GACTGGAGAT GCTATGTTCC GCACACAAAT CCGTGCTCTT CTTCGTGCGT | 2160 |

CTGTTACCGG TCAATTGCGT ATCATGTTCC CAATGGTTGC GCTCTTGAAA GAATTCCTG 2220
CAGCGAAAGC AGTCTTTGAT GAAGAAAAAG CAAACCTTCT TGCTGAAGGT GTTGCAAGTG 2280
CGGATAACAT CCAAGTTGGT ATCATGATCG AGATTCCTGC AGCGGCTATG CTTGCAGACC 2340
AATTTGCTAA AGAAGTTGAC TTCTTCTCAA TTGGTACAAA CGACTTGATC CAATATACAA 2400
TGGCAGCAGA CCGTATGAAC GAACAAGTTT CATACCTTTA CCAACCATAC AACCCATCAA 2460
TCCTACGCTT GATTAACAAT GTGATCAAAG CAGCTCACGC TGAAGGTAAA TGGGCTGGTA 2520
TGTGTGGTGA GATGGCTGGT GACCAACAAG CTGTTCCACT TCTTGTCGGA ATGGGCTTGG 2580
ATGAGTTCTC TATGTCAGCA ACATCTGTAC TTCGTACACG CAGCTTGATG AAGAAACTCG 2640
ACACAGCTAA GATGAAGAG TACGCAAACC GTGCCCTTAC AGAATGCTCA ACAATGGAAG 2700
AAGTTCTTGA ACTTCAAAA GAATACGTTA ATTTTGATTA ATCGAAAAGT CCCTGCAACT 2760
CAGTTACAGG GATTTTTTTG ATATTTTAAA AAGAATTTTC AAGAAAATCT TTCTTATAGA 2820
AAGTCCAACC TTGAAAAAGT AGTGGTCAGA AAAAAAATA CTAAATGGT TCATAAAATT 2880
CTTGACAAGT TGGATATTTA GGAGTAAACT ATTAACCAGT TAAGTAATAG AGAGGAGTTT 2940
CTGCAATTTA GAAATGAATT GCAACTAGAA ATATCAAATA GAAAGAGAGT TTCGATGAAA 3000
ATTAATAAGA AATACCTTGT TGGTCTGCG GCACTTTGAT TTTAAGTGTT TGTCTTACG 3060
AGTTGGGACT GTATCAAGCT AGAACGGTTA AGGAAAATAA TCGTGTTCCT TATATAGATG 3120
GAAAACAAGC GACGCAAAA ACGGAGAATT TGACTCCTGA TGAGGTTAGC AAGCGTGAAG 3180
GAATCAATGC TGAGCAAATC GTCATCAAGA TAACAGACCA AGGCTATGTC ACTTCACATG 3240
GCGACCACTA TCATTATTAC AATGGTAAGG TTCCTTATGA CGCTATCATC AGTGAAGAAT 3300
TACTCATGAA AGATCCAAAC TATAAGCTAA AAGATGAGGA TATTGTTAAT GAGGTCAAGG 3360
GTGGATATGT TATCAAGGTA GATGGAAAAT ACTATGTTTA CCTTAAGGAT GCTGCCCCAG 3420
CGGATAACGT CCGTACAAA GAGGAAATCA ATCGACAAA ACAAGAGCAT AGTCAACATC 3480
GTGAAGGTGG AACTCCAAGA AACGATGGTG CTGTTGCCTT GGCACGTTTCG CAAGGACGCT 3540
ATACTACAGA TGATGGTTAT ATCTTTAATG CTTCTGATAT CATAGAGGAT ACTGGTGATG 3600
CTTATATCGT TCCTCATGGA GATCATTACC ATTACATTCC TAAGAATGAG TTATCAGCTA 3660
GCGAGTTGGC TGCTGCAGAA GCCTTCCTAT CTGGTCGAGG AAATCTGTCA AATTCAAGAA 3720
CCTATCGCCG ACAAATAGC GATAACACTT CAAGAACAAA CTGGGTACCT TCTGTAAGCA 3780
ATCCAGGAAC TACAAATACT AACACAAGCA ACAACAGCAA CACTAACAGT CAAGCAAGTC 3840
AAAGTAATGA CATTGATAGT CTCTTGAAAC AGCTCTACAA ACTGCCTTTG AGTCAACGAC 3900

730

| | |
|--|------|
| ATGTAGAATC TGATGGCCTT GTCTTTGATC CAGCACAAAT CACAAGTCGA ACAGCTAGAG | 3960 |
| GTGTTGCAGT GCCACACGGA GATCATTACC ACTTCATCCC TTA CTCTCAA ATGTCTGAAT | 4020 |
| TGGAAGAACG AATCGCTCGT ATTATTCCTT TTCGTTATCG TTCAAACCAT TGGGTACCAG | 4080 |
| ATTCAAGGCC AGAACACCA AGTCCACAAC CGACTCCGGA ACCTAGTCCA GGCCCGCAAC | 4140 |
| CTGCACCAAA TCTTAAAATA GACTCAAAT CTCTTTGGT TAGTCAGCTG GTACGAAAAG | 4200 |
| TTGGGGAAGG ATATGTATTC GAAGAAAAGG GCATCTCTCG TTATGCTCTT GCGAAAGATT | 4260 |
| TACCATCTGA AACTGTTAAA AATCTTGAAA GCAAGTTATC AAAACAAGAG AGTGTTCAC | 4320 |
| ACACTTTAAC TGCTAAAAA GAAAATGTTG CTCCTCGTGA CCAAGAATTT TATGATAAAG | 4380 |
| CATATAATCT GTTAACTGAG GCTCATAAAG CCTTGTTTGA AAATAAGGGT CGTAATCTCTG | 4440 |
| ATTTCCAAGC CTTAGACAAA TTATTAGAAC GCTTGAATGA TGAATCGACT AATAAGAAA | 4500 |
| AATTGGTAGA TGATTTATTG GCATTCCTAG CACCAATTAC CCATCCAGAG CGACTTGGCA | 4560 |
| AACCAAATTC TCAAATTGAG TATACTGAAG ACGAAGTTCG TATTGCTCAA TTAGCTGATA | 4620 |
| AGTATACAAC GTCAGATGGT TACATTTTGG ATGAACATGA TATAATCAGT GATGAAGGAG | 4680 |
| ATGCATATGT AACGCCTCAT ATGGGCCATA GTCACTGGAT TGGAAAAGAT AGCCTTTCTG | 4740 |
| ATAAGGAAAA AGTTGCAGCT CAAGCCTATA CTAAAGAAAA AGGTATCCTA CCTCCATCTC | 4800 |
| CAGACGCAGA TGTTAAAGCA AATCCAAC TGAGATAGTGC AGCAGCTATT TACAATCGTG | 4860 |
| TGAAAGGGGA AAAACGAATT CCACTCGTTC GACTTCCATA TATGGTTGAG CATACAGTTG | 4920 |
| AGGTTAAAA CGGTAATTTG ATTATTCCTC ATAAGGATCA TTACCATAAT ATTAAATTTG | 4980 |
| CTTGGTTTGA TGATCACACA TACAAAGCTC CAAATGGCTA TACCTTGGA GATTGTGTTG | 5040 |
| CGACGATTAA GACTACGTA GAACACCCTG ACGAACGTCC ACATTCTAAT GATGGATGGG | 5100 |
| GCAATGCCAG TGAGCATGTG TTAGGCAAGA AAGACCACAG TGAAGATCCA AATAAGAACT | 5160 |
| TCAAAGCGGA TGAAGAGCCA GTAGAGGAAA CACCTGCTGA GCCAGAAGTC CCTCAAGTAG | 5220 |
| AGACTGAAAA AGTAGAAGCC CAACTCAAAG AAGCAGAAGT TTTGCTTGCG AAAGTAACGG | 5280 |
| ATTCTAGTCT GAAAGCCAAT GCAACAGAAA CTCTAGCTGG TTTACGAAAT AATTGACTC | 5340 |
| TTCAAATTAT GGATAACAAT AGTATCATGG CAGAAGCAGA AAAATTACTT GCGTTGTTAA | 5400 |
| AAGGAAGTAA TCCTTCATCT GTAAGTAAGG AAAAAATAAA CTAATGAAAA ATGAAAGTCT | 5460 |
| CGATAAAGAG GCTTTCATTT TTATTATGTA TATATGTAAA ATTCTTGACA AGCAATATTA | 5520 |
| AAAAGAGTAA ACTATTAACT AGTTAATTAA CCGGTTTATT ACTTTATAGT GAATCAAATA | 5580 |
| TACTTAAGAA AAGAGGAAAG AATGAAAT AATAAAAAAT ATCTAGCAGG TTCAGTGGCA | 5640 |
| GTCTTGCCC TAAGTGTGTTG TTCCTATGAA CTTGGTCGTC ACCAAGCTGG TCAGGTTAAG | 5700 |

731

| | |
|---|------|
| AAAGAGTCTA ATCGAGTTkC TTATATAGAT GGTGATCAGG CTGGTCAAAA GGCAGAAAAC | 5760 |
| TTGACACCAG ATGAAGTCAG TAAGAGGGAG GGGATCAACG CCGAACAAAT CGTCATCAAG | 5820 |
| ATTACGGATC AAGGTATGTG GACCTCTCAT GGAGACCATT ATCATTACTA TAATGGCAAG | 5880 |
| GTCCCTTATG ATGCCATCAT CAGTGAAGAG CTCCTCATGA AAGATCCGAA TTATCAGTTG | 5940 |
| AAGGATTCAG ACATTGTCAA TGAAATCAAG GGTGGTTATG TTATCAAGGT AGATGGAAAA | 6000 |
| TACTATGTTT ACCTTAAGGA TGCAGCTCAT GCGGATAATA TTCGGACAAA AGAAGAGATT | 6060 |
| AAACGTCAGA AGCAGGAACA CAGTCATAAT CACGGGGGTG GTTCTAACGA TCAAGCAGTA | 6120 |
| GTTGCAGCCA GAGCCCAAGG ACGCTATACA ACGGATGATG GTTATATCTT CAATGCATCT | 6180 |
| GATATCATTG AGGACACGGG TGATGCTTAT ATCGTTCCTC ACGGCGACCA TTACCATTAC | 6240 |
| ATTCTTAAGA ATGAGTTATC AGCTAGCGAG TTAGCTGCTG CAGAAGCCTA TTGGAATGGG | 6300 |
| AAGCAGGGAT CTCGTCCTTC TTCAAGTTCT AGTTATAATG CAAATCCAGC TCAACCAAGA | 6360 |
| TTGTCAGAGA ACCACAATCT GACTGCTACT CCAACTTATC ATCAAAATCA AGGGGAAAAC | 6420 |
| ATTTCAAGCC TTTTACGTGA ATTGTATGCT AAACCCCTAT CAGAACGCCA TGTGGAATCT | 6480 |
| GATGGCCTTA TTTTCGACCC AGCGCAAATC ACAAGTCGAA CCGCCAGAGG TGTAGCTGTC | 6540 |
| CCTCATGGTA ACCATTACCA CTTTATCCCT TATGAACAAA TGTCTGAATT GGAAAAACGA | 6600 |
| ATTGCTCGTA TTATTCCCTT TCGTTATCGT TCAAACCATT GGGTACCAGA TTCAAGACCA | 6660 |
| GAACAACCAA GTCCACAATC GACTCCGGAA CCTAGTCCAA GTCCGCAACC TGCACCAAAT | 6720 |
| CCTCAACCAG CTCCAAGCAA TCCAATTGAT GAGAAATTGG TCAAAGAAGC TGTTCGAAAA | 6780 |
| GTAGGCGATG GTTATGTCTT TGAGGAGAAT GGAGTTTCTC GTTATATCCC AGCCAAGGAT | 6840 |
| CTTTCAGCAG AAACAGCAGC AGGCATTGAT AGCAAACTGG CCAAGCAGGA AAGTTTATCT | 6900 |
| CATAAGCTAG GAGCTAAGAA AACTGACCTC CCATCTAGTG ATCGAGAATT TTACAATAAG | 6960 |
| GCTTATGACT TACTAGCAAG AATTCAACCA GATTTACTTG ATAATAAAGG TCGACAAGTT | 7020 |
| GATTTTGAGG CTTTGGATAA CCTGTTGGAA CGACTCAAGG ATGTCyCAAG TGATAAAGTC | 7080 |
| AAGTTAGTGG ATGATATTCT TGCCTTCTTA GCTCCGATTC GTCATCCAGA ACGTTTAGGA | 7140 |
| AAACCAAATG CGCAAATTAC CTACACTGAT GATGAGATTC AAGTAGCCAA GTTGGCAGGC | 7200 |
| AAGTACACAA CAGAAGACGG TTATATCTTT GATCCTCGTG ATATAACCAG TGATGAGGGG | 7260 |
| GATGCCTATG TAACTCCACA TATGACCCAT AGCCACTGGA TTAAAAAAGA TAGTTTGTCT | 7320 |
| GAAGCTGAGA GAGCGGCAGC CCAGGCTTAT GCTAAAGAGA AAGGTTTGAC CCCTCCTTCG | 7380 |
| ACAGACCATC AGGATTCAGG AAATACTGAG GCAAAAGGAG CAGAAGCTAT CTACAACCGC | 7440 |

732

| | |
|---|------|
| GTGAAAGCAG CTAAGAAGGT GCCACTTGAT CGTATGCCTT ACAATCTTCA ATATACTGTA | 7500 |
| GAAGTCAAAA ACGGTAGTTT AATCATACCT CATTATGACC ATTACCATAA CATCAAATTT | 7560 |
| GAGTGGTTTG ACGAAGGCCT TTATGAGGCA CCTAAGGGGT ATACTCTTGA GGATCTTTTG | 7620 |
| GCGACTGTCA AGTACTATGT CGAACATCCA AACGAACGTC CGCATTCAGA TAATGGTTTT | 7680 |
| GGTAACGCTA GCGACCATGT TCGTAAAAAT AAGGTAGACC AAGACAGTAA ACCTGATGAA | 7740 |
| GATAAGGAAC ATGATGAAGT AAGTGAGCCA ACTCACCTG AATCTGATGA AAAAGAGAAT | 7800 |
| CACGCTGGTT TAAATCCTTC AGCAGATAAT CTTTATAAAC CAAGCACTGA TACGGAAGAG | 7860 |
| ACAGAGGAAG AAGCTGAAGA TACCACAGAT GAGGCTGAAA TTCCTCAAGT AGAGAATTC | 7920 |
| GTTATTAACG CTAAGATAGC AGATGCGGAG GCCTTGCTAG AAAAAGTAAC AGATCCTAGT | 7980 |
| ATTAGACAAA ATGCTATGGA GACATTGACT GGTCTAAAAA GTAGTCTTCT TCTCGGAACG | 8040 |
| AAAGATAATA ACACTATTTT AGCAGAAGTA GATAGTCTCT TGGCTTTGTT AAAAGAAAGT | 8100 |
| CAACCGGCTC CTATACAGTA GTAAAAAGT TGGAGCATAT TTTATGGAGA AGTAACCTTT | 8160 |
| CGTGTTACTT CTCTTTTTTA GAAAAACGTA ACAGA | 8195 |

(2) INFORMATION FOR SEQ ID NO: 95:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2004 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 95:

| | |
|--|-----|
| TTTACTAAAA GGAAAAAGA ACTGATTCTT CAGTCCTTCA TTAATCTTAT TCCACACTAA | 60 |
| ATAGGTATGG GTAAACAGGT TGTGACCTT GGTGAATCTC GACTTCAACG TCTTCGAATT | 120 |
| CTTCTACGAT TTCTTGAGCG ATTTTCATTGG CAAGTTCTTC GCTTCCGTCT TCACCTACAT | 180 |
| AGAAGGTTAC GATTTCACTG TCTTCATCCA ACATATGTTT CAAGGTTTCA GTCAATGTTT | 240 |
| GGTGCATATC AGGGTTTGAC ACAAGAATTT TTCCATCCAC CATACCTAAA TTATCGTTTT | 300 |
| CATGGATTTC TAAGCCATCG ATCGTTGTAT CACGCACGGC TGTGTGACG CTTCGGCTAA | 360 |
| CGACATCGCT AAGAGCAGCT GTCATACGCT CTTGGTTTTT TTCAATGGAC TTGCTTGGAT | 420 |
| CAAAGGCAAG AAGACTTGTC ATACCTTGAG GAAGAGTGCG AGCCTCTACC ACTACCGCTG | 480 |
| GTTGCTCCAA AACTTCTGCC GCAGATTGAG CTGCCATGAA GATGTTCTTG TTGTTTGGCA | 540 |
| AGAAGATGAT GTTACGGGCA TTAACCTGTT CAACAGCCTT GATAAAGTCT TCTGTTGAAG | 600 |
| GGTTCATGGT TTGACCGCCT TCGATAACAT AATCCACGCC TTGAGAACAG AAGATATCTG | 660 |

733

| | |
|---|------|
| CTAGACCTTT ACCAGCCACC ACAGCAATCA AAGCATACTC TTTTCTTCA GCCGACTTGA | 720 |
| TAACCTGAGT AGCTTCTTTC TCAACCTGTG CTTCTGTGTG GTTACGCATA TTGTCAACTT | 780 |
| TTACCTTGAC CAAGCTACCA TATTTGAGAC CTTCTTGCAT AACAAAGTCCT GGATCTTCTG | 840 |
| TATGAACATG GACTTTGACA ATTTTCATCAT CGTTAACAAC AAGGAGAGAA TCTCCAAGCT | 900 |
| CATCCAAGTA GTTACGGAAT TCATCGTAGT CAAAATCTTT AGCATAGGTT GGACCTTGCT | 960 |
| TAAGAGCTAC CATGATTTCA GTACAGTAAC CAAACGTGAT GTCCTCAGTC GCTACGTGAC | 1020 |
| CAGCTACAGA CTTATGATGC TCTACATTGA TCATCTCACT CATGTTGGCA GGAGTCGCTA | 1080 |
| CAAAGTCCTC AGATGCAATA TATTCGCCAG TAAGGGCTGA AAGGAAACCT TCGTAGATGA | 1140 |
| AGACCAATCC TTGACCACCT GAGTCCACAA CGCCAACTTC TTTCAATACT GGAAGCATGT | 1200 |
| CTGGTGT TTT AGCTAGAGCT GTTTTAGCAC CTTCCAAGGC TGC GCGCATG ACTTCAACAG | 1260 |
| CGTCATCTGT TTGCTCAGCT TTTTCTTAG CACCGATAGC AGCTCCACGA GAAACTGTTA | 1320 |
| AAATCGTTCC TTCAACAGGT TTCATCACTG CCTTATAGGC AACTTCCACA CCTGATTGGA | 1380 |
| AGGCCAGAGC CAAGTCTTGA CCTGTAACT CGTCTTATC CTTGATAGCT TGGGAAAATC | 1440 |
| CACGGAAG CTGAGACGTA ATCACTCCTG AGTCCACG CGCACCACATC AAAAGCCCTT | 1500 |
| TGGCAAGAAT GCTCGCTACT TCTCCAACG TAGAAGCTGG CTTGTCTGCA ACTTCTTAG | 1560 |
| CACCATTTTC AATGGTCATT CCCATATTTG TCCCAGTATC TCCATCTGGA ACTGGAAGA | 1620 |
| CGTTTAATGA ATTGACATAT TCAGCTTGCT TATTCAGCG AGTTGATGCA GCCTGCACCA | 1680 |
| TTTCTTGAAA TAAGCTAGTA GTAATTTT ACACGGTTAT TCTCTACAA CTTTGATATT | 1740 |
| TTGAATGTAG ACATTTACAG TCTGAGCAGT AATTCCAAGC TGGTTTTCCA AGCTAAAGGC | 1800 |
| AACACGCTCT TGAATGTTTT TTGACACTTC ACTAATCTTT GTTCCGTAGC TTAACACGGT | 1860 |
| ATATACATCA ACTGCAATAC TGCCATCTTC GGCTGCCTTT ACGACGACAC CTTTAGAATA | 1920 |
| ATTTTCCTTA CCTAGCAGGG CTTGGAAATT ATCTTTGAGG GCATTTTAC TAGCCATACC | 1980 |
| GACCACACCA GAAATCTCAG TTGC | 2004 |

(2) INFORMATION FOR SEQ ID NO: 96:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 11915 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 96:

734

| | |
|---|------|
| CCGGGTGGG CTGTTGCCCC ATTAAGCGG CACCACAGCT GGGTTCAGAA CGTCGTGAGA | 60 |
| CAGTTCGGTC CCTATCCGTC GCGGGCGTAG GAAATTTGAG AGGATCTGCT CCTAGTACGA | 120 |
| GAGGACCAGA GTGGACTTAC CGCTGGTGTA CCAGTTGTCT TGCCAAAGGC ATCGCTGGGT | 180 |
| AGCTATGTAG GGAAGGGATA AACGCTGAAA GCATCTAAGT GTGAAACCCA CCTCAAGATG | 240 |
| AGATTTCCCA TGATTATATA TCAGTAAGAG CCCTGAGAGA TGATCAGGTA GATAGGTTAG | 300 |
| AAGTGAAGT GTGGCGACAC ATGTAGCGGA CTAATACTAA TAGCTCGAGG ACTTATCCAA | 360 |
| AGTAACTGAG AATATGAAAG CGAACGGTTT TCTTAAATTG AATAGATATT CAATTTTGAG | 420 |
| TAGGTATTAC TCAGAGTTAA GTGACGATAG CCTAGGAGAT ACACCTGTAC CCATGCCGAA | 480 |
| CACAGAAGTT AAGCCCTAGA ACGCCGGAAG TAGTTGGGGG TTGCCCCCTG TGAGATAGGG | 540 |
| AAGTCGCTTA GCTCTAGGGA GTTTAGCTCA GCTGGGAGAG CATCTGCCTT ACAAGCAGAG | 600 |
| GGTCAGCGGT TCGATCCCGT TAACTCCCAT TTTAGCGGGT GTAGTTTAGT GGTAAACTA | 660 |
| CAGCCTTCCA AGCTGTTGTC GCGAGTTCGA TTCTCGTCAC CCGCTTTGAA CTTTGTCTTT | 720 |
| TGTACCAAGT TTTTGACTTG GCGCGTAGC TCAGGTGGTT AGAGCGCACG CCTGATAAGC | 780 |
| GTGAGGTCGG TGGTTCGAGT CCACTCGTGC CCATAGTGT TAGTCCATTA CTAGGGGATT | 840 |
| GGAATATTAT CTGTTCACTA AGAGGACACG GGCTTGTTCC CGTATAAACT ATTTTGAGG | 900 |
| ATTACCCAAG TCCGGCTGAA GGGAACGGTC TTGAAAACCG TCAGGCGTGT AAAAGCGTGC | 960 |
| GTGGGTTCGA ATCCACATC CTCCTTTTAT ATTAACGCGG GATGGAGCAG CTCGGTAGCT | 1020 |
| CGTCGGGCTC ATAACCGAA GGTGCTAGGT TCAAATCCTG CTCCCGCAAT AAGGCTCGGT | 1080 |
| AGCTCAGTTG GTAGAGCAAT GGATTGAAGC TCCATGTGTC GGCGGTTGTA TTCCGTCTCG | 1140 |
| CGCCATTTAT ATATTTGGA AGGGTAGCGA AGAGGCTAAA CGCGCGGAC TGTAATCCG | 1200 |
| CTCCTTCGGG TTCGGGGGTT CGAATCCCTC CCCTTCCATT TTACGGGCAT AGTTTAAAGG | 1260 |
| TAGAACTAAG GTCTCCAAAA CCTTCAGTGT GGGTTCAATT CCTACTGCCC GTGTTAATAG | 1320 |
| AATTATGGCG GGTGTGGTGA AGTGGTTAAC ACACCAGATT GTGGCTCTGG CATGCGTGGG | 1380 |
| TTGATCCCC ATCACTCGCC TATTTTATAT TGGGGTATAG CCAAGCGGTA AGGCAAGGGA | 1440 |
| CTTTGACTCC CTCATGCGTT GGTTCGAATC CAGCTACCCC AGTTACTATT TGCCGGCGTG | 1500 |
| GCGGAATTGG CAGACGCGCT GGAATCAAAA TCCAGTGTCC GCAAGGACGT GCCGGTTCGA | 1560 |
| CCCCGGCCGC CCGTATAGTA TAGTGTTAGG AACGTTGTTA TTCTTCGTTT CTTTTTATA | 1620 |
| TTATTTTGG TATAATTATA GTTATTCAAA TTTTATTTAG ATTAAGAAAG TGTAGGGGAG | 1680 |
| TATGTCPTGT TCTATCGATT TATTAAACA TCGGTATTTG AAAAATATTA AAGAAAATCC | 1740 |
| TGAATTGTTT GTCGAATTG AGTTGGAGTA TCCTGTTGCA AGTTTAGAAG GGGATGCTAC | 1800 |

735

| | |
|--|------|
| AGATGTTGAA GTTATGAAGG ATCTATTTCA TTATTTAGTT TCTACTTTGG ATCTCACCGT | 1860 |
| AGCAAAGGTA GATGATTTTG GCAATCTGAT CCAGTTAGTA GATCCGATAA GTCAGGATGC | 1920 |
| TATTTTATTT GAAGTTTCCT ATACAACGAT TGAGTTTGCA TTTGGTAAGG CTGAAACGAT | 1980 |
| TCAAGAGGTC GAAAATCGTT TCAATAATTA TATGAATGTA ATTCAGAGAA AGTTAGCTGA | 2040 |
| ATCAAATCAT GCTATGTG GCTGTGGTAT CCATCCCAAC TGGGATAAAA ATGAGAATTG | 2100 |
| TCCAGTGGCT TATCCACGCT ATCAGATGTT GATGGATTAT TTGAATTTGA GTAGAAATAT | 2160 |
| TATTAAATCA GATTTACATC ATTTCCCTGA ATATGGTACT TTTATCTGTG GGAGCCAGGT | 2220 |
| TCAGCTGGAT ATTTCAAAAA CCAACTACTT ACGGGTGATT AATGCTTTTA CTCAAATTGA | 2280 |
| AGCGGCTAAG GCTTATTTAT TTGCAAACTC TGAATTTTCG GGTGCGGATT GGGATACGAA | 2340 |
| AATTTCAAGG GATATTTTCT GGGAAGAATC TATGCATGGT ATCTATCCAG AGAATGTTGG | 2400 |
| GGTCAATGCT AGACTCCTTA ATGATGAAAC TGATTTTFTT GACTATCTAA ATCATTCTGC | 2460 |
| GATTTTFACT GCGGAACGTG ATGGGCAGAC CTATTATTTT TATCCTATTC AGGCTGGGGA | 2520 |
| CTATTTGGCT ACGTCCGAAA TCCAAGCATT TGCTCTGAAT GGGGATGAGG TTATTATTTA | 2580 |
| CCCCAAGAG AAGGATTTTG AAATCATCG TAGTTACCAG TACCAAGATT TAACGACTCG | 2640 |
| AGGAACAGTT GAGTTTCGTA GTGTGTGTAC ACAGCCACTT GATAGGACTT TTGCTTCTGC | 2700 |
| AGCTTTTCAC TTGGGATTAT TGGTTAATTT AGACAAGTTA GAAGCTTACT TAGAAACAGC | 2760 |
| ACCTTTCTTT AAAGTATTTG GTTATGATTA CAACTCTTTA AGGAGACAAT TTTCTAAGAA | 2820 |
| AAATCTTACA GATGAGGAAG AAATACGAT TATTGAATTT TCCAAAGACT TACTCCTACT | 2880 |
| AGCTGAGGAG GGACTAGTGG TGAGAAATAA GGAAGAAATG ACCTATTTAC AGCCTTTGAG | 2940 |
| AGAAGAATTG AGCCTATAAT TTCTCTTATA AAGGGAGAAT TTTCTGAAAA ATCATGATAT | 3000 |
| AATGGACGAG ACTATAGATA AAGGATAGAG AGTAATGACA TTAGTTTATC AATCAACGCG | 3060 |
| TGATGCCAAC AATACAGTAA CTGCCAGCCA AGCAATTTTG CAAGGTTTGG CGACGGACGG | 3120 |
| CGGTTTGT TT ACACCGGATA CTTATCCAAA GGTAGATTTG AACTTTGACA AATTGAAAGA | 3180 |
| TGCTTCTTAC CAGGAAGTTG CTAAGCTAGT TTTGTCAGCA TTTTATAGATG ACTTTACAGT | 3240 |
| TGAGGAGTTG GACTACTGTA TCAACAATGC CTACGATAGC AAATTTGATA CTCCAGCTAT | 3300 |
| TGCACCATTA GTGAAATTAG ATGGGCAATA CAATTTGGAA CTTTCCATG GTTCAACGAT | 3360 |
| TGCCTTTAAG GATATGGCCT TGTCTATTTT GCCATACTTT ATGACGACTG CTGCTAAGAA | 3420 |
| ACATGGTTTG GAGAACAGA TTGTTATCTT GACAGCGACA TCTGGTGACA CGGGGAAAGC | 3480 |
| TGCTATGGCG GGGTTTGC GA ATGTGCCTGG TACTGAGATT ATCGTCTTTT ATCCAAAGGA | 3540 |

736

| | |
|--|------|
| TGGTGTGAGC AAGATTCAAG AGTTACAAAT GACCACTCAG ACTGGCGACA ATACTCATGT | 3600 |
| TATTGCTATT GATGGTAACT TTGACGATGC GCAAACAAAT GTGAAGCACA TGTTTAACGA | 3660 |
| CGTGGCTCTT CGTGAAAAAT TGACTACCAA CAAGTTGCAA TTTTCATCAG CTAACCTAT | 3720 |
| GAACATTGGT CGTCTGGTGC CACAAATTGT TTATTATGTT TATGCTTACG CTCAATTGGT | 3780 |
| TAAGACTGGT GAAATTGTAG CTGGTGAAAA GGTAACTTC ACAGTACCAA CAGGAACTT | 3840 |
| TGGAAATATC TTGGCTGCCT TTTATGCCAA ACAAATTGGT TTGCCAGTTG GTAAATTAAT | 3900 |
| CTGTGCTTCA AATGACAACA ATGTTTGAC AGACTTCTTT AAAACACGTG TCTATGACAA | 3960 |
| AAAACGTGAG TTAAAGGTAA CAACCAGCCC ATCTATGGAT ATCTTGGTAT CTTCAAACCT | 4020 |
| GGAGCGCTTG ATTTTCCATC TTTTGGGAAA TAATGCTGAA AAGACAACGT AACTTATGAA | 4080 |
| TGCCTTGAAC ACGCAAGGAC AATATAAGTT GACAGACTTT GATGCAGAGA TTTTGGACCT | 4140 |
| CTTTGCAGCT GAATATGCGA CTGAGGAAGA AACGGCAGCA GAGATCAAGC GTGTTTGTGA | 4200 |
| GTTAGATTCT TATATCGAGG ACCCTCATAC AGCTGTTGCT TCAGCAGTTT ATAAAAATA | 4260 |
| CCAATCGGCC ACTGGAGATG TAACTAAGAC AGTGATTGCT TCAACAGCTA GTCCATACAA | 4320 |
| GTTCCAGTA GTTGCAGTAG AAGCTGTAAC TGGAAAAGCA GGTTTAACAG ACTTTGAAGC | 4380 |
| CTTGGCTCAA TTACATGAAA TCTCAGGCGT TGCAGTGCCA CCAGCAGTTG ATGGGCTTGA | 4440 |
| AATAGCTCCA ATTCGTCACA AGACAACAGT GGCAGCTGCT GACATGCAAG CAGCGGTTGA | 4500 |
| GGCTTATTTA GGACTTTAAG ACAGAGGGAG CAAACTCGGT TGGGAAACCA ACTGAGTTTC | 4560 |
| TTTTTCATCAG GAGGAGAGAT TGTTTAAGAA AAATAAAGAC ATTCTTAATA TTGCATTGCC | 4620 |
| AGCTATGGGT GAAAACTTT TGCAGATGCT AATGGGAATG GTGGACAGTT ATTTGGTTGC | 4680 |
| TCATTTAGGA TTGATAGCTA TTTCAGGGGT TTCAGTAGCT GGTAATATTA TCACCATTTA | 4740 |
| TCAGGCGATT TTCATCGCTC TGGGAGCTGC TATTTCCAGT GTTATTTCAA AAAGCATAGG | 4800 |
| GCAGAAAGAC CAGTCGAAGT TGGCCTATCA TGTGACTGAG GCGTTGAAGA TTACCTTACT | 4860 |
| ATTAAGTTTC CTTTtaggat TTTTGTCCAT CTTGCTGGG AAAGAGATGA TAGGACTTTT | 4920 |
| GGGGACGGAG AGGGATGTAG CTGAGAGTGG TGGACTGTAT CTATCTTTGG TAGGCGGATC | 4980 |
| GATTGTTCTC TTAGGTTTAA TGACTAGTCT AGGAGCCTTG ATTCGTGCAA CGCATAATCC | 5040 |
| ACGTCTGCCT CTCTATGTTA GTTTTTTATC CAATGCCTTG AATATTCTTT TTTCAAGTCT | 5100 |
| AGCTATTTTT GTTCTGGATA TGGGGATAGC TGGTGTGCT TGGGGACAA TTGTGTCTCG | 5160 |
| TTTGGTTGGT CTTGTGATTT TGTGGTCACA ATTAAACTG CCTTATGGGA AGCCAACTTT | 5220 |
| TGGTTTAGAT AAGGAACTGT TGACCTTGGC TTTACCAGCA GCTGGAGAGC GACTTATGAT | 5280 |
| GAGGGCTGGA GATGTAGTGA TCATTGCCCTT GGTGCTTTCT TTTGGGACGG AGGCAGTTGC | 5340 |

737

| | |
|--|------|
| TGGGAATGCA ATCGGAGAAG TCTTGACCCA GTTTAACTAT ATGCCTGCCT TTGGCGTCGC | 5400 |
| TACGGCAACG GTCATGCTGT TGGCCCGAGC AGTTGGAGAG GATGATTGGA AAAGAGTTGC | 5460 |
| TAGTTTGAGT AAACAAACCT TTTGGCTTTC TCTGTTCCCTC ATGTTGCCCC TGTCCTTTAG | 5520 |
| TATATATGTC TTGGGTGTAC CATTAACCTCA TCTCTATACG ACTGATTCTC TAGCGGTGGA | 5580 |
| GGCTAGTGTT CTAGTGACAC TGTTTTCACT ACTTGGGACC CCTATGACGA CAGGAACAGT | 5640 |
| CATCTATACG GCAGTCTGGC AGGGATTAGG AAATGCACGC CTCCCTTTTT ATGCGACAAG | 5700 |
| TATAGGAATG TGGTGTATCC GCATTGGGAC AGGATATCTG ATGGGGATTG TGCTTGGTTG | 5760 |
| GGGCTTGCCCT GGTATTGGG CAGGGTCTCT CTTGGATAAT GGTTTTCGCT GGTATTCTT | 5820 |
| ACGCTATCGT TACCAGCGCT ATATGAGCTT GAAAGGATAG GAAATGCAAA AAACAGCTTT | 5880 |
| TATTTGGGAT TTAGACGGGA CTTTATTGGA CTCTTACGAA GCGATTTTAT CAGGGATTGA | 5940 |
| GGAGACTTTT GCTCAGTTT CTATTCCCTA TGATAAGGAG AAGGTGAGAG AGTTTATCTT | 6000 |
| CAAGTATTCG GTGCAAGATT TGCTTGTCG GGTGGCAGAA GATAGAAATC TGGATGTTGA | 6060 |
| GGTGCTAAAT CAGGTGCGTG CCCAGAGTCT GGCTGAGAAG AATGCTCAGG TAGTTTGTAT | 6120 |
| GCCAGGTGCG CGTGAGGTGC TAGCTTGGGC AGACGAATCA GGAATTCAGC AGTTTATATA | 6180 |
| TACTCATAAG GGAACAACG CTTTACCCT TCTCAAGGAC TTGGGGGTGG AATCCTATTT | 6240 |
| TACAGAGATT TTAACCAGTC AGAGTGGCTT TGTGCGGAG CCAAGTCCAG AAGCGGCTAC | 6300 |
| CTATCTGCTA GATAAGTATC AGTTGAATTC TGATAATACT TATTATATAG GGGATCGGAC | 6360 |
| TCTGGATGTG GAATTTGCCC AGAATAGTGG GATTCAAAGC ATCAACTTTT TAGAGTCTAC | 6420 |
| TTATGAAGGG AATCACAGGA TTCAAGCGTT AGCAGATATT TCCCGTATTT TTGAGACTAA | 6480 |
| GTGATAAAAA GATTGTGTCA GTTTTGTGAC AGAGACCTAA CAACTATTT CAAGTAACCT | 6540 |
| AGTTTGTAC AAGGAATAGA CAGTTCTGTT AAATAGGCC GAGAGGGCTT TTTTCTACA | 6600 |
| TTTTTTGTGT TATGATAGAC AGGTACTCAT TTGAAAGGAA TTTGAAAGAA TGAAGAAAAG | 6660 |
| AATGTTATTA GCGTCAACAG TAGCCTTGTC ATTTGCCCA GTATTGGCAA CTCAAGCAGA | 6720 |
| AGAAGTTCTT TGGACTGCAC GTAGTGTGA GCAAATCCAA AACGATTGA CTAACACGGA | 6780 |
| CAACAAAACA AGTTATACCG TACAGTATGG TGATACTTTG AGCACCATTG CAGAAGCCTT | 6840 |
| GGGTGTAGAT GTCACAGTGC TTGCGAATCT GAACAAAATC ACTAATATGG ACTTGATTTT | 6900 |
| CCCAGAACT GTTTTGACAA CGACTGTCAA TGAAGCAGAA GAAGTAACAG AAGTTGAAAT | 6960 |
| CCAAACACCT CAAGCAGACT CTAGTGAAGA AGTGACAACT GCGACAGCAG ATTTGACCAC | 7020 |
| TAATCAAGTG ACCGTTGATG ATCAAACGT TCAGGTTGCA GACCTTTCTC AACCAATTGC | 7080 |

738

| | |
|---|------|
| AGAAGTTACA AAGACAGTGA TTGCTTCTGA AGAAGTGGCA CCATCTACGG GCACTTCTGT | 7140 |
| CCCAGAGGAG CAAACGACCG AAACAACTCG CCCAGTTGAA GAAGCAATC CTCAGGAAAC | 7200 |
| GACTCCAGCT GAGAAGCAGG AAACACAAGC AAGCCCTCAA GCTGCATCAG CAGTGGAAAT | 7260 |
| AACTACAACA AGTTCAGAAG CAAAAGAAGT AGCATCATCA AATGGAGCTA CAGCAGCAGT | 7320 |
| TTCTACTTAT CAACCAGAAG AGACGAAAAT AATTTCAACA ACTTACGAGG CTCCAGCTGC | 7380 |
| GCCCGATTAT GCTGGACTTG CAGTAGCAAA ATCTGAAAAT GCAGGTCTTC AACCACAAAC | 7440 |
| AGCTGCCTTT AAAGAAGAAA TTGCTAACTT GTTTGGCATT ACATCCTTTA GTGGTTATCG | 7500 |
| TCCAGGAGAC AGTGGAGATC ACGGAAAAGG TTTGGCTATC GACTTTATGG TACCAGAACG | 7560 |
| TTCAGAATTA GGGGATAAGA TTGCGGAATA TGCTATTCAA AATATGGCCA GCCGTGGCAT | 7620 |
| TAGTTACATC ATCTGGAAAC AACGTTTCTA TGCTCCATTC GATAGCAAAT ATGGGCCAGC | 7680 |
| TAACACTTGG AACCCAATGC CAGACCGTGG TAGTGTGACA GAAAATCACT ATGATCACGT | 7740 |
| TCACGTTTCA ATGAATGGAT AAACCCGACT TGATAACATC ATTTTGACGA ATGAGATCTA | 7800 |
| GCTTTCGTGA TGGAAAGCGA TTCTCGTTCG TTTTTCCTTT GTCATACTCT TCGAAAATCT | 7860 |
| CTTCAAACCA CGTCAGTTTT ATCTGAACT TCAAAGCTGT GCTTTGAGCA ACCTGCGACT | 7920 |
| AGCTTCCTAG TTTGCTTTTT GATTTTCATT GAGTATCAAT TTGAATGGAA AATGGAAAGT | 7980 |
| TATCATCTTG TAATGAGTTA AGCAACATTC TTGCAATCTA TTTTACTTTA TATCACAATT | 8040 |
| AATTAGTCAA ATATTGATAA ATCAATAAAA AGAGAGGGGA AGAAATGCTA GAGATTCAAG | 8100 |
| ATTTACTGTA TCAACTCCGC TTGTCTGAGC AAGCGAGTAC GCAATTGTTT GAAAAAGGC | 8160 |
| TTGGGATTAG TTTGACACGG TATCAGATTT TACTGTTTTT GCTGGAGCAT TCTCCTTGTA | 8220 |
| ACCAAATGGC GGTTCAGGAG CGTTTGAAAA TTGATCAGGC TGCTTTGACA CGGCATTTCA | 8280 |
| AAATTTTGA AACGGAAGGT TTGGTGGAGC GTCATCGTAA TCCTGAAAAT CAGCGGGAAG | 8340 |
| TGTTGGTAGA GGCTGCGAAG TATGCCAAGG AGCAGTTAGT GGTGAATCCC CCTCTGCAAC | 8400 |
| ATATCAGGGT TAAGGAAGAG ATAGAAAGTA TCTTAACAGA GTTTGAGAGA ACAGAACTCA | 8460 |
| GCCGTTTATT AAATAAATTG GTTTTGGGTA TTGAAAATAT AGAAATTAA GGAGAAATAG | 8520 |
| ATGTCAATTA TTTTAACAAC GATCGTTGCT TTGGAGCATT TTTACATTTT TTATTGGAA | 8580 |
| AGTATTGCCA CGCAATCAGA TGCGACTAGT CGTGTATTTA ATATGGAAAA GGAAGAATTG | 8640 |
| GCTCATCCGT CAGTAAGTTC ATTGTTCAAA AATCAAGGAA TTTATAAGGC TCTGCTAGGA | 8700 |
| GTCTTCTCT TGTATGTCAT TTATTTCTCA CAGAAATTAG AAATTGTGAC TATTTTGTG | 8760 |
| TTATTTGTGA TTGGTGCTGC GACTTACGGC TCTTTAACAG CGGATAAAAA AATTATTTTG | 8820 |
| AAACAAGGTG GATCAGCTAT TTTGGCCTTG ATTAGTATTT TACTCTTTAA ATACACTTGA | 8880 |

| | |
|--|-------|
| AGGTCGATTC TAATCTCGCT AATCCTTTTT AATCCAGAAT AAGGGAAATA TGTTATACTT | 8940 |
| GTTTTTAAGA AAAAAGTCTC ATTGAATTGG TTTTGAGGAG TTAGAAATGA AAGTATTAGT | 9000 |
| GACAGGTTTT GAGCCCTTTG GAGGGGAAAA GGGCAATCCA GCTTTGGAGG CCATTAAAGG | 9060 |
| TTTACCAGCT GAAATCCATG GTGCTGAGGT CCGTTGGCTA GAGGTGCCGA CAGTTTTTCA | 9120 |
| CAATCTGCT CAAGTATTGG AAGAAGAGAT GAATCGTTAT CAACCTGACT TTGTCTTTG | 9180 |
| TATTGGGCAA GCTGGTGAA GAACTAGTTT GACACCTGAA CGAGTGACCA TTAATCAAGA | 9240 |
| CGATGCATGC ATTTCTGATA ACGAAGATAA TCAACCGATT GACCGTCCCA TTCGCCCAGA | 9300 |
| TGGTGCTTCG GCCTACTTTA GTAGTTTGCC GATTAAAGCG ATGGTCAAG CTATAAAAAA | 9360 |
| AGAGGGCTTA CCGGCCTCTG TTTCCAATAC GGCAGGGACT TTTGTCTGCA GCCATTTGAT | 9420 |
| GTATCAGGCT CTCTATTTGG TAGAAAAGAA ATCTCCATAT GTTAAGGCAG GTTTTATGCA | 9480 |
| TATTCCTTAT ATGATGGAAC AGGTGGTGAA CAGACCGACT ACTCCAGCTA TGAGTTTAGT | 9540 |
| GGATATTCGG CGAGGGATAG AAGCAGCAAT CGGCGCTATA ATAGAACATG GAGATCAGGA | 9600 |
| ACTCAAGTTG GTAGGCGGAG AAATCATTG ATAGAAAAA GCTTGAGGGG AAAAACCTTC | 9660 |
| AAGCTTTTGG ACGTTTTTGG GCCAATACTG CTCGGTAAAA CATAATTTTA GTGCATTGGA | 9720 |
| TATAAGGTAG GAGTGAAAAA CTAGCAATGC CAAAGGTAAT CCAATTGAGG AAGTACCAAG | 9780 |
| GAAGAAGCTG TAAATCTAGG ACAAAGTGCT GGAAGTTGTA GCCCTTCATA AAGGAACGGC | 9840 |
| TAGTTTTTAG GATTCGTCTT GGTGGGACCT GTCCTAGGTC TAGACTATAA CAGAGAAGAA | 9900 |
| ATTCCACCTG TGAATAGGCA TAATACTGTG GAATATAGAG GATATTTCTT ACAATGATCA | 9960 |
| AGATGAGACT TGCAAGAAAG TAGAGTCCAA AGACCATGAG GAAACGCTCG GTTTCAACTG | 10020 |
| ATGAGAGATC TAGATTTGGA AACTCAGGAT GTAGGGTGAC GAATTTTTTG GCTAAAAAGC | 10080 |
| TACTATAAAA GAGGAGGTAA ATCCCAAGTA AATTAGGGAT ACTCCATAAA AAGAGATAGA | 10140 |
| AACGTTTGAG AAGTAGGGTC AAAAAGGTTT GAGAAAAGCG CTCCTCATCA AAGAGAGCTA | 10200 |
| GGCTGTTTTT TACAGATGGC TCCGTTTTAG AATCTTTCAT GAGTGTCAGT GTTGATAGA | 10260 |
| CGGAACTGGT CAAAAGAATA GTCCCGATAA AGGAGACTAG TAGAGGAAAG AGGTAGGTTT | 10320 |
| GAAGTATTTG GCCAAGTATG CTGAAAAATG GCTGTTCTAA AACAGTCCCG TGGATCCGAG | 10380 |
| ATAAGGGATT AAGAAAACCA GATAAGATGA CCAGCATACT GGGAAAGGATA TAGAGGAGAA | 10440 |
| AGAGACGGGG GGTGTCAGCC TGAAAATGTT TTGACTCCTG ACGAATTGTT TTTAAATCAA | 10500 |
| TTTTTGGATA GTTCATTCTC TTATTATACC ATAGTTCTTA TACATAGTTC GTGACAGTTC | 10560 |
| CTACTTTTTT TGATAAAATC ATACAGTGTG TCCTTGGGCA CACTGTATGA ACTGGGACTG | 10620 |

740

| | |
|---|-------|
| TCTTTCCCAG CTTCCGAGGT AAAAAATGTC AGATTCACCA ATCAAATATC GTTTGATTAA | 10680 |
| GAAAGAAAA CACACAGGAG CTCGTCTGGG AGAAATCATC ACTCCCCACG GTACCTTTPCC | 10740 |
| GACACCTATG TTTATGCCAG TTGGGACACA AGCCACTGTC AAAACTCAGT CACCTGAAGA | 10800 |
| ATTGAAGGAG ATGGGTTCGG GAATTATCCT ATCAAACACC TATCATCTCT GGCTTCGCCC | 10860 |
| TGGAGATGAA CTCATTGCAC GCGCTGGTGG TCTCCACAAG TTCATGAATT GGGACCAGCC | 10920 |
| TATCTTGACA GATAGTGGTG GTTTTCAGGT TTATTCTTTA GCAGATAGCC GTAATATCAC | 10980 |
| AGAAGAAGGA GTAACCTTTA AAAATCATCT AAATGGTTCT AAGATGTTCC TATCCCCAGA | 11040 |
| AAAAGCCATC TCTATTGACA ATAATCTGGG TTCAGACATC ATGATGTCCT TTGATGAATG | 11100 |
| TCCTCAGTTT TATCAACCTT ATGACTACGT TAAGAAATCG ATCGAGCGTA CCAGCCGTTG | 11160 |
| GGCTGAGCGT GGTTTGAAGG CTCACCGTCG TCCACATGAC CAAGGTTTGT TTGGAATTGT | 11220 |
| GCAAGGTGCA GGATTTGAAG ACCTTCGCCC CCAATCAGCT CATGATCTTG TCAGCATGGA | 11280 |
| TTTCTCAGGC TACTCTATCG GTGGTTTGGC AGTGGGAGAA ACCCATGAAG AGATGAATGC | 11340 |
| GGTCTTGAC TTTACAACCT AACTGCTGCC TGAATAATAA CCTCGTTATC TGATGGGTGT | 11400 |
| GGGAGCGCCA GATAGCTTGA TCGATGGGGT CATTCGTGGG GTGGATATGT TTGACTGTGT | 11460 |
| CTTACCGACT CGAATTGCTC GTAACGGGAC TTGTATGACC AGTCAAGGAC GTTTGGTTGT | 11520 |
| GAAAAATGCC CAGTTTGCTG AGGACTTTAC GCCACTGGAT CCTGAGTGTG ATTGCTACAC | 11580 |
| ATGTAATAAC TATACACGCG CTTACCTTCG TCACCTGCTC AAGGCTGATG AAACCTTTGG | 11640 |
| TATCCGCTTG ACTAGCTACC ACAATCTTTA CTTCTTGCTT AACCTGATGA AGCAAGTGCG | 11700 |
| ACAAGCCATC ATGGATGACA ATCTCTTGA ATTCCGTGAG TATTTTGTGG AAAAATATGG | 11760 |
| CTATAATAAG TCAGGACGTA ATTTCTAAAA TGAATTGAT ATAAAAAAT CCTAAGTTT | 11820 |
| CTCTTAGGAT TTTTCTTCTT TTTTGTATAG AATAAAGTGT ACAATGAAAG GAAGAATAAA | 11880 |
| CTCGTATGCG CATTAAATGG TTTTCCTCGA TTAGG | 11915 |

(2) INFORMATION FOR SEQ ID NO: 97:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 9069 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 97:

| | |
|--|-----|
| GAGAGGGCAA CAGTTCTATC GCTTCAAAT TTTTCTTGGT TTGCAGATAT TCAAGAATCG | 60 |
| GGAGTTTTC TATAGTATTC GGCAGATTTA TTACAGCCAA GCATCTCAA AATACGGACA | 120 |

741

| | |
|---|------|
| GCATCCTCCA TCTTTTCTG GCCTTCCTTG ACTCTACCTT GCTTGCTATC AAGGAGACCT | 180 |
| TCTGCCACCA GATAACAAT TCGGAAATAG GTCTCATTTC CTTGTAGAA ATGCTCTTCG | 240 |
| ATAACACGTT TAAATAATA GGCATTGGTA AATTCTTCAC ACTCAATACT AGCTAAAAAG | 300 |
| CCATTCAATA GTATAGTATG AAAAAGGTTT CGATTGCCAG ACATTTCAT TAGAAAATCA | 360 |
| GATTTACGTA CCATTTCTCG TACATATCTA GTAAAAAGAG AAACAGATAA AAATGGAGAA | 420 |
| CTGACTGAAA ATAAATTGAG TTCATAGATT CCCAGATCT CGGTAGAAAA CAAATAATCA | 480 |
| TGAAGGACTT TTCCTTCCTC TGCTGTTAAG TCTACCCTTT CATCTATGCT CTTCATATAA | 540 |
| GACTTGATAA TAATGGCATT TAGAATATGT TTCTGTTTGT TGTGAGAATG GGCATGCTTT | 600 |
| TATACTCCCT GCGATATAAG TCCTCAAGAG GTGCTATATT CTTTGGTTCC AAGACATCTG | 660 |
| TAATTTCTTT TCTCAACTCA GAATCTGTAT CATACTGGAA ACCTCTTGCC AGAAAGAGGA | 720 |
| TCTCCTCCAC ACTGGCAGAT ATATTTTCCA GAGCAAATAG AAATTTTCC ACCGAAAGCT | 780 |
| CACTCTGACC TGTTCAAAA CGGGACAACA TAGACGGCGA AAATTGTCCT CCGGTTGCTT | 840 |
| GTCTCAGTGA GATATTTCTT GACTCTCGTA ATTGTCTAAA GACTTTTCCA ATCTGCTCCA | 900 |
| TAGACTTCCC CTTGATTCCG TATTTTCTTC ATTTTATCAT ATTTTTCAGA AAATTCATCA | 960 |
| AAACTTGCC AAATTGTCAG AATTATGAGA AAATAGAGGA TATTTATCAC GTGGAGGGAC | 1020 |
| TGCTATGAGA GACGATATCA AAATCAATGA CCGTGCTTTG GCCTTGCAAG ACCAAATTAT | 1080 |
| CGAAAACTA GAGAAAGTTT TTGATACAGA TGTGGAATTG GATGTTTACA ATCTAGGTCT | 1140 |
| GATTATGAA ATCAATCTGG ATGAAACGGG GCTCTGCAAG ATTGTCATGA CCTTCACCGA | 1200 |
| TACTGCCTGT GATTGCGCCG AAAGCCTGCC TATTGAAATC GTGGCAGGTC TGAAACAAAT | 1260 |
| CGAGGTATC AAAGATATCA AGGTTGAAGT TACCTGGTCG CCGCTTGGA AAATCACACG | 1320 |
| AATCAGTCGC TATGGCCGTA TTGCCCTTGG ACTACCACCT CGTTAAGCAG ACCAATCACT | 1380 |
| TTTAAAGATG AAAATCAAAG GGCAAACTAG AAAACTAGCC GCAGGTGCT CAAAACACTG | 1440 |
| TTTTGAAGTT ATGGATAGAA CTGACGAAGT CAGCTCAAAA CACTGTTTTG AGGTTGTGGA | 1500 |
| TAGAACTGAC GAAGTCAGCT CAAAACACTG TTTTGAGGTT GTGGATAGAA CTGACGAAGT | 1560 |
| CAGCCCAAAA CACTGTTTTG AGGTTGTGGA TAGAACTGAC GAAGTCAGTA ACCATACCTA | 1620 |
| CGGCAAGGCG ACGTTGACGT GATTGAAGA GATTTTCGAG TATGAGTTTA TTTTTACCT | 1680 |
| GACTTGTTCA TATTCAGAA GTCTGTCACG GCTCCGCGTG AAGCAGATGA TACGATGTGG | 1740 |
| GCATATTTAC CGAGGACACC ACGGCTGTAA AGTGGTGGCA AGGTTGTTTC TGCCTTGCGT | 1800 |
| TTTTCAAGTT CTCTTCGGA TACGGCCATA GAAATTTCTT TGGTATCTTG GTCAACCGTA | 1860 |

742

| | |
|--|------|
| ACGATATCGC CGGTACGGAG ATAGGCAATT GGTCCACCAT CCTGAGCTTC AGGAGCGATA | 1920 |
| TGTCCAACAA CCAGACCATA AGTACCACCA GAGAAACGTC CGTCCGTCAA GAGGGCCACC | 1980 |
| TTATCTCCCT GACCTTTACC AACAAATCATT GAAGAAAGTG ATAGCATCTC AGGCATACCA | 2040 |
| GGACCACCTT TAGGTCCAAC AAAACGAACA ACGACTACAT CGCCATCAAC GATTTTCATCT | 2100 |
| GTCAGAACGG CCTGAATCGC ATCTTCTTCT GAGTCAAAGA CCTTAGCTGG CCCAACGTGA | 2160 |
| CGACGCACTT TAACACCTGA TACCTTGGCA ACTGCACCGT CAGGAGCAAG GTTCCCCTTC | 2220 |
| AAGATGATAA GCGGACCATC CGCACGTTT GATTTTCAA GTGGCATGAT AACTTTTGG | 2280 |
| CCTGGAGTCA AGTCTGCAA GTCAGCCAAG TTTTCAGCTA CAGTCTTACC AGTACATGTG | 2340 |
| ATGCGATCTC CGTGAAGGAA ACCATTTGCC AACAAATACT TCATAACCGC AGGGACACCA | 2400 |
| CCGACTTCGT AGAGGTCTTG GAAGACATAC TGACCAGATG GTTTCAAGTC GGCCAAGTGA | 2460 |
| GGCACACGTT CTTGAATCGT ATTGAAGTCC TCAAGTGACA AGTCAACATT TGCGGCATGG | 2520 |
| GCAATGGCGA GCAAGTGAAG AGTGGCGTT GTAGAACCAC CGAGAGCCAT CGTTACAGTG | 2580 |
| ATAGCATCTT CAAAGGCTTC ACGAGTCAAG ATATCTGATG GTTTGAGACC AAGTTCCAAC | 2640 |
| ATCTTAACAA CAGCACGTCC TGCTGCTTCG ATATCTTCTT TCTTATCAGC TGATTACAGT | 2700 |
| GGGTGAGAGG ATGACCCTGG CAAACTCATC CCTAGAACTT CGATAGCAGT TGCCATGGTA | 2760 |
| TTAGCAGTAT ACATACCACC ACAACCACCA GGGCCAGGGC AGGCATTACA TTCAAGACGT | 2820 |
| TTACAGTCCT CAGCTGTCAT GTCACCGTGG TTCCATTTTC CGATACCTTC AAAGACAGAA | 2880 |
| ACCAAGTCGA TATCTTTACC ATCAAGATTT CCCGGTGCAA TAGTTCCACC ATAGGCGAAA | 2940 |
| ATAGCTGGGA TATCCATATT AGCAATAGCA ATCATAGATC CAGGCATGTT CTTGTACACAG | 3000 |
| CCACCGATAG CGACGAAGGC ATCCACGTTG TGACCACTCA TAGCCGCCTC GATGGAGTCC | 3060 |
| GCGATGATGT CACGAGATGT TAGAGAGAAA CGCATACCAG GCGTTCCCAT AGCGATCCCG | 3120 |
| TCCGCTACGG TAATGGTTCC AAAGTGTACA GGCCAAGCGC CTGCAGATTT GACACCTTCT | 3180 |
| TTAGCCAGTT TCCCGAAATC ATGCAAGTGA ATGTTACATG GTGTATTTTC CGCCCAAGTC | 3240 |
| GAAATCACTC CCACAATCGA TGTTTCAAAG TCCTTATCTG TCATACCAGT CGCACGAAGC | 3300 |
| ATAGCACGGT TAGGTGATTT AACCATGCTG TCATAAATGC TACTGCGGTG ACGTTTATCT | 3360 |
| AATTCAGTCA TCTTATCCCT CCCATTTCAG TTTTACTAT TATAGCACAA TTTTCGCATG | 3420 |
| AAGAACAGAA TAAAATCTT GAATTTTCAG AAAATTCTAT ACACATGTGA AATATTTAAA | 3480 |
| ATTA AAAACA ACAAGCGGA TTAGTGCACT TTCTGATGAC CAGAATATGC TTTTAAATCC | 3540 |
| GCTTTCTTTA AATAACGTAC TGTAATTTT ACAGAAATTC TTTCAAATAA GTGTATTTAA | 3600 |
| CATCTATCTT GCATTATAAA TTTCTAGAAC CTCTCTTTT ATATTCGATT CACTCAAACC | 3660 |

743

| | |
|--|------|
| ATACTCATTA AGAAGATAAT CCATTTTCCC TACTTGACCG AATCTTTCTT GAACACCCAT | 3720 |
| CCGATGAATT TTTGTTATTC CATCATCAGA GAATAATTCA CATAAAGCAC TGCCAATTCC | 3780 |
| ACCTATCTGA TTGTGGTTTT CTACAGTAAA TATAGTTTTT CCACTTAACA TTGTTTTTAT | 3840 |
| CTGTTCTGGT ATCGGTTTGA TTCTAAATAA ATCTATCACA CCTACTGAAT AACCTAATTT | 3900 |
| AGACAGTTCA TCTGCAACTC GAATACTTGG AGCAACCATT ATGCCAGAAG CAACGATTAC | 3960 |
| AAGATCTTCA CCATGCCTTA ACTCAATGTA GCCTTTAGAA AAATCTTCTC CACCTTGATA | 4020 |
| CACAGGAACT GGAGCTTTTC TAATTGTTTCG AATATATTTT AGTCCTTTTA AGTCTAATGT | 4080 |
| CTGGTTCAAT ATTTACAGAA ATTGGATATC ATCAGTTGCT TCGAAAATGA TTGATTTAGG | 4140 |
| AATTAAACGT AACAATCCAA TTTCTTCAAA TGGCATATGT GTTCCACCAT TCATCTCTGC | 4200 |
| CGTTACTCCT GCATCTGATC CAATCACAGT GGCATCCAAT TGTGCGTATC CAAGAGAAAT | 4260 |
| AAATAATTGA TCAAATACTC TTCGTGAAGC AAAAGGACCA AATGTATGAA GATAAGGTCT | 4320 |
| AAACCCCTGA ATAGACAAGC CTGCTGCAAG GCCGACCATT TCTGCTTCCA TAATCCCAAC | 4380 |
| ATTACATAA CGGTCTCCAA AGTCCTTTTC AAGATTATTA GTAGCCATCG AACTTGACAA | 4440 |
| ATCGGCTTCT AAGACTACTA TATCAGAATC ACTTTGATTA GCCTCTAAAA GGAAGTCTCT | 4500 |
| ATATACATGC CGTAATTCTT TCGTACTTCT CATCATTTCTG TTTCTCTCAA TTCCTGACTT | 4560 |
| AATCTTTCTA CAACTGAAGT TAACATTTGT TTCTCCTCTA CAGTAGGGCG AAGATGATGA | 4620 |
| TTGGATTTC A TTTCTCCAG CTCTTGAACC CCTTGACCTT TAATAGTATC TAATACAATG | 4680 |
| CACTTAGGTG ATGAATTATT TGACTGTTTT AATTGGACAA TCCCTTCATA AATTTCTCTA | 4740 |
| ATATCTGAAC CCTTGACCCT AATGGATTCA AATCCAAATG CTGAAAATTT TTCTACGAAA | 4800 |
| TCACCTGGAT TACAAATATC CTTTGTAATA CCATCTAATT GTTTTTTGTT ATCATCAACA | 4860 |
| AATACAATTA AGTTGGATAA CTGTTGATGA GAAGCAAACCT GTATAGCCTC CCAACATTGT | 4920 |
| CCCTCATTTA ACTCACCATC TCCAACAATA GCGTAAGTAT AAAAGGGACT CTTTCTTATT | 4980 |
| CTCTGACCAT ATGCAAGTCC AGTTGCAACA CTAATTCCTT GTCCTAAAGA GCCCGTTGTC | 5040 |
| ATATCTATGC CTGGCGTTAG ATTTCTATCA GGATGAGACG GTAATTGGT TCCATTTGTA | 5100 |
| TTTAAAGAAT ATAAGAATTC TTGTCAAAG AAACCATTCA AATAGAGTGT ACTGTATAGA | 5160 |
| GCTGGTCCTC CGTGACCTTT TGATAATATG AAATAATCTC TATCTCGTGC TGCAAATATT | 5220 |
| TCTGGAGTCA TTGGCATTAT TTCACCATAA AGCACCGCTA AAACCTCTAC GATAGACAGA | 5280 |
| CTTCCTCCGT AATGTCCGAA TCCAAGATGA TTCAATGTTT TAAGAGTATT TAATCGGATG | 5340 |
| TTAGTCGCAA ATTTTCTTAA CCCATCTTCT CTATTTTAC TTAAAAATCAT CCCTTATTCC | 5400 |

744

| | |
|---|------|
| TCCGTTGCAG ATGGCTTTT AATAAAGGAT ACTCCAAACA TAACTGCTAG AATAAGAACA | 5460 |
| AGACCAATCA CAATGCCTGC TTGTGAGCCA AATTGATTTA ACATTCTAA AATAATTCCT | 5520 |
| GATAGACCAA AATCTGCATC TGAGAAAGTT GATCCTTGA AACCAAGTCC TCCCAAACT | 5580 |
| GGCATTAAAA AGACTGGAAG AAAACTGATT AAAATACCTT GTAAAAATGC TCCAATAGTG | 5640 |
| GCTCCACGAA CACCACCAGA TGCATTCCCA ATGACACCTG CAGTCGCTCC ACAGAAGAAA | 5700 |
| TGAGGCACAA CACCTGGTAA GATAACAACC GTTCCTGAAG CAATCATAAT TACCATACTT | 5760 |
| ACTAAACCAC CAACAAACT AGAGATAAAT CCAATTAGAA CTGCATTGGG TGCATAAGTA | 5820 |
| TAAACAATCG GACAATCCAA AGCAGTTTT GAATTAGGTA CAAGACGCTC TGAAATACCT | 5880 |
| TTAAAGGCTG GAACAATTC GCCCAAAATA AGGCGAACAC CTGCTAAAAT AACAAATACC | 5940 |
| CCTGCTGCAA ATTGACCTGC TAATTGTAAA GCATAAACTA GACCACTTGT ACCACTACTG | 6000 |
| ATTTCTTTT CTATATATTC TGACCCTGCA AAGATAGCTA CAATAATGTA AATAACTGCC | 6060 |
| ATGGATAAAG TAATACTAAC AGTACTATCA CGTAAAAAG CTAAACTCTT TGGAAATTTA | 6120 |
| ATGTCCTCTG TTGATTTGA TTTGTCACCG ATAAGGCTAC CAGTAAAACC ACTCAACCAA | 6180 |
| TATCCCAAAG AACTGAAATG ACCTAAAGCT ACCTTGTCAT TTCCAGTTAA TTGAACCATA | 6240 |
| TATTTTGTGA CAAATGCTGG GGAAATACTC ATAATAATAC CGAGTGCTAA TCCTCCTAGT | 6300 |
| AAGATGAGAG GCAAGCTAGT AAAGCCAGCA ACTGATAAAA TGACCGCAAT CATACTGCC | 6360 |
| ATATATAGAG TGTGGTGCCC TGTAAAAAA ATATATTTAA ATCGAGTAAA ACGAGCGATT | 6420 |
| AAGATATGTA ACACCATGCC TGCAAACATA ATCATTGCAG TAGCTGAGCC ATATGTTGTT | 6480 |
| AAAGCTACAG CTACAATTGC TTCATTATTC GGCACAACGC CAGATAAATG AAAAGCATGC | 6540 |
| TCAAACATGG TACCAAATGG ATTTAAAGAA TTTGTACAA TTCCTGCACC ACCAGATACA | 6600 |
| ACTAAGAAAC CAACAAAGT CTTAATTCCA CCTTTAATAA TATCAGGTAA TTTCTTCTTC | 6660 |
| TGAAGAACTA ATCCTAAGAT TGCAATTAAA GCTACTAAAA TAGCTGGTGT ACTAACAATA | 6720 |
| TCCAATATGA ACTTCATCAT GACGCTAGCC TCCTATATAA GTCCTTTTTC TTCACAAAGT | 6780 |
| TTAGTAATTA ATTCTCGTAG TTCATCCATA TCAATAATAC TATTTAAGAT ACGAACATCT | 6840 |
| CCAAGATGAC TAGCTGAATC AGCTAGATCA CGACCAACAA TCCAAATATC AGCTGCATTT | 6900 |
| GGATCTGCTC CACCTAAATC ATAATGTTCA ACTTCTACAT CCGAAACATT CAAATCACTC | 6960 |
| AATACAGATT CAATATTCAT CTGTACCATA AACTTGAAC CTAATCCTGA ACCACAAGCT | 7020 |
| GTACCAATTT TTAACATTAT CTAATCCTCC TGTTTAATTA TCATTTTAAT GTCATCATAG | 7080 |
| TTTTTTGATG ATATTAAAGT TTGAACATGA TTTTATCTC TAAAATTGT TGTAAATGT | 7140 |
| GACAAAGCCT TTAAATGACT CTCATTATCA ATGGCTGCAA TACAAATCAA CAATCTTACC | 7200 |

745

| | |
|--|------|
| TCTTGTCTG GATTATCCAA TAAATAAATC GGTTCCTCCA AAAC TAACAT TGACATTCCT | 7260 |
| ATTTCATTCA CACCTTCATC TGGCCGAGCG TGAGGAATTG CTA CTCCCTT CCCTAAATTA | 7320 |
| ATAAAAGGTC CAAACTCTTC TACTTTTGA ATCATTGCCT CAGGGTAGTT CTCAGTTATC | 7380 |
| TTATCTTGAT CCAAAGCGG TTTAGCTGCT AAACGAATCG CCTCCTCCA TCCTAATTTT | 7440 |
| TGCGAACTAA CTGATAGGT TTCTTTGGTA ATAAGTTGTT CTAGCACTGG TACAATTTCC | 7500 |
| TTTCTATCAT TTTTTTGGTA AAGATAATTC TTAAACGCCA ATCTTAATTC CAATTCCTGT | 7560 |
| GTAATAATTC CATATCTTTT GACAATATTC AGGATTGTT CAATCTCAA ATCTCCATAC | 7620 |
| TCTAAATTCG GAAAACTTTT TAACACTAGT TCTACTAGTT GTATTGCTTG CTCTTCAGTC | 7680 |
| ATCATAACCG AAAC TAGATA ATTTGGCTTT TCTGTCTCCA CCTTTATGGT AGAAAAAACC | 7740 |
| ATATCATAGT CACTACTAGC TTTCACCTGT AAATCATCAA TCTTTGAGGT TCCTATAAAC | 7800 |
| TCAATTTGAG GAAATAATGC TAATAGATTC TCTTTTAACA TCAATGAAGA ACTAACACCA | 7860 |
| TTAGGACAAA TGATTGCTGC TTTATACCAT TTTTGAGGCA AAGTATCTGC TTTCTTTAAA | 7920 |
| TAACCTCCGA AATGGATAAC AAAATATGCT GTTTCACTAT CAGGTATGGG ATTGTCAATA | 7980 |
| GCGTCCATCA AGGGCATCAA AGAATCTTTG ACTAATTCOA ATAAATCAGG ATAATGTTCT | 8040 |
| TTAACATGCA ATACATATTC ATTTGAACTA GGTAGGCCGA ACTTTAATCT ATAGTAAGCC | 8100 |
| GGTATAAGGT GCGGCGAAG ATTTCTCTC AATCCTTCCC TTGTGTTAAA ATGTAACAAA | 8160 |
| GAAATATCTT CCATTCTACT TATAATAGCC TCTGTTAATT GATTAAAGTA AACC GGAGCA | 8220 |
| ACATCTACTT CACCTTCAA GCAACTTGAT AATAAAACGG TGATATAGCG ATAATCATCC | 8280 |
| TCAGAAAACA CCGTATCTAT AATTCCTAAA TCAACCACTG TATCCAATAA AATAGTGGTT | 8340 |
| ATATCTTGAA TAACAGGAGA TACTAATGTC TCTGAAAGAC ATACTCTTTC AACATCCCTT | 8400 |
| TGATACCTAC ACAGAATGAA TACTAAACCG AAAAGGTAAA CTTTTAATTG ATTAACAATA | 8460 |
| GGTACTAGCT GTAGCTTCTC ATAATAATCT TTAACCTACCT GATCAATCAA ATCATAAGTT | 8520 |
| AATGAATACC CCCAACTGGA TAAAACATAA TCCAAACCCC AAATCCCTAT GGAGGATTCC | 8580 |
| AGCAACTCAC TAACCATTG AAAAGCTAAG CGGTGCTTAT TCCACTCTGA ACCGTGTAAA | 8640 |
| GTATAACCTT TTGCTCTACT GTACCCTAGC TCCAAATCAT TATCTAACAT AATCTTTCTT | 8700 |
| AATGATTGAA TATCAGATAA GGTGTATTC TTA CT TACTT TCAAAAAGTC TTGGTAATGA | 8760 |
| CTATTGATA TAAAATCTAA TCGGCAAAAA GTGTAAAGAT AGATTAAAGC TAAGCGAGTC | 8820 |
| GACTTTGGTA AAACCAATTC ATCCGACTTA ATAATATCTG TCAAAGACTG CTTCTGACGA | 8880 |
| TTTGATAAAC TATAGCGACC TTGCTTTTTA TCCAGCACTA TCCCTTTATT AGCTAGATAA | 8940 |

746

| | |
|---|------|
| GGCACTAAAT AATCTATTCC TTCTTTGACT TCCTTTATAG GTAAGCTCAC CTTAACAGAT | 9000 |
| AATTCATATA ACGATAGCTC ACAATGATCC ATCAAAGTCA TCAAAATAAC TAGTGCTCTA | 9060 |
| TAATCAAAC | 9069 |

(2) INFORMATION FOR SEQ ID NO: 98:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 8654 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 98:

| | |
|--|------|
| CGAGACAACA AGATGAAGAA AAATTTGCCC TATCGTTTGT GCGCCTTGCA AGTGTAGCAC | 60 |
| TTCTTGCAGC CTGTGGAGAA GTGAAGTCTG GAGCAGTCAA CACTGCTGGT AACTCAGTAG | 120 |
| AGGAAAAGAC AATTAAATC GGGTTTAACT TTGAAGAATC AGGTTCTTTA GCTGCATACG | 180 |
| GAACAGCTGA ACAAAAAGGT GCCCAATTGG CTGTTGATGA AATCAATGCC GCAGTGGTAT | 240 |
| CGATGGAAAA CAAATCGAAG TAGTCGATAA AGATAATAAG TCTGAAACAG CTGAGGCTGC | 300 |
| TTCACTTACA ACTAACCTTG TAACCCAATC TAAAGTATCA GCAGTCGTAG GACCTGCGAC | 360 |
| ATCTGGTGGC ACTGCAGCTG CGGTAGCGAA CGCTACAAAA GCAGGTGTTC CATTGATCTC | 420 |
| ACCAAGTGCG ACTCAAGATG GATTGACTAA AGGTCAAGAT TACCTCTTTA TTGGAACCTT | 480 |
| CCAAGATAGC TTCCAAGGAA AAATTATCTC AAAGTATGTT TCTGAAAAAT TAAATGCTAA | 540 |
| GAAAGTTGTT CTTTACACTG ACAATGCCAG TGAATATGCT AAAGGGATTG CAAAATCTTT | 600 |
| CCGCGAGTCA TACAAGGGTG AAATCGTTGC AGATGAAACT TTCGTAGCAG GTGACACAGA | 660 |
| CTTCCAAGCA GCCCTTACAA AAATGAAAGG GAAAGACTTT GATGCTATCG TTGTTCCCTGG | 720 |
| TTACTATAAT GAGGCTGGTA AAATTGTAAA CCAAGCGCGT GGCATGGGAA TTGACAAACC | 780 |
| AATCGTTGGT GGTGATGGAT TCAACGGTGA GGAGTTTGTA CAACAAGCAA CTGCTGAAAA | 840 |
| AGCATCAAAC ATCTACTTTA TCTCAGGCTT CTCAACTACT GTAGAAGTTT CAGCTAAAGC | 900 |
| TAAAGCCTTC CTTGACGCTT ACCGTGCTAA GTACAATGAA GAGCCTTCAA CATTTGCAGC | 960 |
| CTTGGCTTAT GATTCACTTC ACCTTGTAGC AAACGCAGCA AAAGGTGCTA AAAATTCAGG | 1020 |
| TGAAATCAAG AATAACCTTG CTAAAACAAA AGATTTTGAA GGTGTAACCTG GTCAAACAAG | 1080 |
| CTTCGATGCA GACCACAACA CAGTCAAAAC TGCTTACATG ATGACCATGA ACAATGGTAA | 1140 |
| AGTTGAAGCA GCAGAAGTTG TAAAACCATA ATAGAAAAAT GTTGAAATAG GGAATGAGCC | 1200 |
| TTTGACTCAC TCCCTGTTTC GATATTTAAT ACTCTTCGAA AATCTCTTCA AACTGCGTCA | 1260 |

747

| | |
|--|------|
| ACGTCGCCCTT GGATTATATA TGTGACTGAC TTCGTCAGTC TTATCTACAA CCTCAAAGCA | 1320 |
| GTGCTTTGAG CAACCTGCGG CTAGTTTCCT AGTTTGCTCT TTGATTTTCA TTGAGTATAA | 1380 |
| GAACCTATCA AAAAGTGAGG GAAAACCCCTC GGAATTATAA ATAGAAAGAG TGAATCTTAT | 1440 |
| GCTCCAACAA CTCGTAAATG GTTTGATTCT AGGTAGTGTT TACGCGCTGT TAGCCCTAGG | 1500 |
| ATATACCATG GTTTACGGAA TTATCAAGCT CATCAACTTC GCCCATGGTG ATATTTATAT | 1560 |
| GATGGGAGCC TTTATCGGTT ATTTCTTGAT CAATTCCTTC CAAATGAATT TCTTTGTAGC | 1620 |
| GCTTATTGTA GCTATGCTAG CGACAGCTAT TCTTGGTGTC GTGATTGAGT TTCTTGCTTA | 1680 |
| CCGACCTTTG CGCCACTCTA CTCGTATTGC TGTTTTGATT ACGGCTATTG GGGTTCTTTT | 1740 |
| CCTATTGGAG TATGGAATGG TCTATCTGGT TGGTGCCAAT ACCCGTGCCT TCCCTCAAGC | 1800 |
| GATTCAAACA GTTCGATATG ATTTGGGACC AATTAGCTTA ACAAATGTGC AGTTAATGAT | 1860 |
| TTTGCCCATT TCCTTGATTT TGATGATTTT GTTACAAGTC ATTGTCCAAA AGACTAAGAT | 1920 |
| GGGGAAAGCC ATGCGTGCCAG TATCAGTAGA TAGCGACGCG GCGCAATTGA TGGGGATCAA | 1980 |
| TGTAAACCGT ACGATTAGCT TTACCTTCGC TTTGGGTCTT GCTCTTGCGG GTGCGGCTGG | 2040 |
| TGTTCTGATT GCTCTTTATT ATAACCTCTT TGAGCCTTTG ATGGGGGTTA CTCCAGGTCT | 2100 |
| TAAATCTTTC GTTGCCGCAG TACTTGGTGG TATCGGAATT ATTCTGGTG CGGCTCTTGG | 2160 |
| TGGCTTTGTG ATTGGTCTAT TGGAAACCTT TGGCACTGCC TTTGGGATGT CAGATTTCCG | 2220 |
| TGATGCCATT GTTTATGGAA TCTTGTGTGTT GATCTTGATT GTCCGCCAG CTGGTATCCT | 2280 |
| TGGTAAGAAT GTGAAAGAGA AGGTGTAAAC GATGAAGGAA AATTATAAAG TTAATATTCT | 2340 |
| ATGGTTACTC CTTTGTGTAG CTGGCTATAG CTTGATTAGT GTACTGGTTT CAGTCGGAGT | 2400 |
| ACTTAATCTA TTCTATGTAC AGATTTTACA ACAAATTGGA ATTAATATTA TTTGGCTGT | 2460 |
| TGGTCTCAAC TTAATCGTTG GTTTTTCAGG ACAATTTTCA CTTGGTCATG CTGGTTTCAT | 2520 |
| GGCGATTGGT GCCTATGCAG CAGCTATTAT TGGTTCTAAA TCACCAACCT ACGGTGCCTT | 2580 |
| CTTTGGAGCT ATGCTTGTAG GGGCTTTGCT TTCAGGAGCA GTTGCCTTAC TTGTCGGCAT | 2640 |
| TCCAACCTTG CGCTTGAAGG GGGACTATCT TCGGCTAGCA ACTCTGGGTG TTTCTGAAAT | 2700 |
| TATCCGTATC TTTATCATCA ATGGTGGAAG CCTTACAAAT GGTGCGGCAG GTATCTTAGG | 2760 |
| GATTCCTAAC TTTACAACCT GGCAAATGGT TTACTTCTTT GTCGTGATTA CAACCATTGC | 2820 |
| AACCTTGAAC TTCTTGCGTA GCCCAATTGG TCGTTCAACC CTCTCTGTTT GTGAAGATGA | 2880 |
| AATCGCTGCT GAGTCAGTTG GGGTTAATAC GACTAAAATT AAAATCATCG CTTTGTCTTT | 2940 |
| TGGTGCCATT ACTGCAAGTA TTGCTGGGTC ACTTCAGGCA GGATTTATCG GGTCTGTTGT | 3000 |

748

| | |
|--|------|
| ACCGAAAGAT TACACCTTCA TCAACTCAAT CAACGTTTTG ATTATTGTTG TATTTGGTGG | 3060 |
| ACTCGGTTC ATTACAGGTG CGATTGTTTC GGCTATTGTT CTGGGAATTT TGAATATGCT | 3120 |
| TCTCCAAGAT GTTGCTAGTG TGCCTATGAT TATTACGCT TTGGCCTTGG TATTGGTAAT | 3180 |
| GATTTTCAGA CCAGGTGGAC TCCTTGGAAC ATGGGAAC TGAGCTATCAC GTTCTTTTAA | 3240 |
| AAAATCTAAG AAGGAGGAAC AAAACTAATG GCATTACTTG AAGTAAACAA GTTAACCAAA | 3300 |
| CATTTTGGTG GTCTAACAGC TGTGGGAGAT GTGACTCTTG AATTGAACGA AGGGGAAC TG | 3360 |
| GTGGATTAA TCGGTCCAAA CGGAGCTGGG AAAACCACCC TTTTCAACCT TTTGACCGGT | 3420 |
| GTTTATGAAC CAAGCGAGGG AACAGTAACC CTAGATGGTC ACCTTTTGAA TGGGAAATCA | 3480 |
| CCTTATAAGA TTGCCTCTTT GGGACTTGGG CGTACTTTCC AAAATATCCG TCTCTTTAAA | 3540 |
| GATTTAACAG TTTTAGATAA TGTTTTGATT GCTTTTGGA ACCATCACAA ACAGCATGTT | 3600 |
| TTTACTAGTT TCTTACGCTT ACCAGCTTTT TACAAGAGTG AAAAAGAATT AAAGGCTAAA | 3660 |
| GCTTTGGAAT TGTGAAAAT CTTTGATTTA GATGGTGATG CAGAGACTCT TGCTAAAAAT | 3720 |
| CTTCTCTACG GACAACAACG TCGTTTGGA ATTGTTTCGT CCCTTGCTAC GGAACCTAAA | 3780 |
| ATTCTCTCT TAGATGAACC AGCAGCAGGT ATGAACCCAC AGGAAACAGC CGAATTGACT | 3840 |
| GAGTTAATTC GTCGTATCAA AGATGAGTTT AAGATTACAA TCATGTTGAT TGAACACGAT | 3900 |
| ATGAATCTGG TCATGGAAGT AACAGAACGT ATCTACGTAC TTGAATATGG CCGTTTAATC | 3960 |
| GCTCAAGGAA CTCCAGACGA AATTAAGACC AATAAACGCG TTATCGAAGC TTATCTAGGA | 4020 |
| GGTGAAGCCT AATGTCTATG TTAAAAGTTG AAAATCTTTC TGTGCATTAC GGTATGATCC | 4080 |
| AAGCAGTTCG TGATGTAAGC TTTGAAGTTA ATGAAGGAGA AGTTGTTTCC CTTATCGGTG | 4140 |
| CCAACGGTGC AGGTAAGACA ACTATTCTTC GCACCTTGTC AGGTTTGTTT CGACCAAGTT | 4200 |
| CAGGAAAGAT TGAATTTTGA GGTCAAGAAA TCCAAAAAAT GCCAGCTCAG AAAATCGTGG | 4260 |
| CAAGTGGTCT TTCACAAGTT CCAGAAGGAC GCCACGCTTT TCCTGGCTTG ACTGTTATGG | 4320 |
| AAAATCTTGA AATGGGAGCT TTCTTAAAGA AAAATCGTGA AGAAAATCAA GCTAACTTGA | 4380 |
| AGAAGGTTTT CTCACGCTTT CCTCGTCTTG AAGAACGGAA GAACCAAGAT GCAGCCACTC | 4440 |
| TTTCAGGGGG GGAACAACAA ATGCTTGCCA TGGGACGCGC CCTCATGTCA ACACCAAAAC | 4500 |
| TTCTTCTTTT AGATGAACCA TCAATGGGAC TTGCCCAAT CTTTATCCAA GAAATTTTGT | 4560 |
| ATATCATTTA AGATATTGAG AAGCAAGGAA CAACGGTCCT CTGTATTGAA CAAAATGCCA | 4620 |
| ATAAAGCACT TGCAATCTCT GACCGAGGAT ATGTACTGGA AACAGGGAGA ATCGTCCTAT | 4680 |
| CAGGAACAGG AAAAGAATCT GCTTCATCAG AAGAAGTCAG AAAAGCATAT CTAGGTGGCT | 4740 |
| AAAACAATCC AGTGGATTGT TTTAGTCGGC AGATGGAGAT TACGAAGTAA TCATCAATAT | 4800 |

749

| | |
|--|------|
| AGTCCGGGGG ACCTTTT TAG TCGGTAGATT GAGATTGCAA ACAAATCTGC ATCTACATTG | 4860 |
| AAAGCTTAAT TTCTAATAAT TGAAAAAATC GAATGAAAAA TTTCTTACCT TCATTCACAG | 4920 |
| AGCTCGATTT CAGAGCTCTT TTTGCTAGCT TATTCATACT TTTCTGAATT TCGAAAAAGA | 4980 |
| AATGTAAGCG TTTGATAGAT TTACAAAAAG ATTGTATAAT AGGGATAAGA ATAGAAAAGG | 5040 |
| AGAAGTCTCA TGGCAGTTAA AGATTTTATG ACCCGCAAGG TAGTTTATAT TAGTCCAGAT | 5100 |
| ATAACAGTAT CTCATGCAGC AGATTTGATG AGAGAGCAAG GTTTGCACCG TCTGCCTGTT | 5160 |
| ATCGAAAAATG ATCAATTAGT TGGTTTGGTG ACTGAGGGAA CCATTGCACA AGCAAGTCCA | 5220 |
| TCTAAAGCAA CAAGTCTTTC TATCTATGAG ATGAATTATC TTCTGAATAA GACAAAAGTA | 5280 |
| AAAGATGTCA TGATTCGCGA TGTGTGCACT GTCTCAGGCT ATGCTAGTCT AGAAGATGCA | 5340 |
| ACTTATCTGA TGTGAAAAA TAAGATTAGT ATTCTCCCTG TCGTAGATAA CCATCAAGTA | 5400 |
| TACGGAGTTA TTACTGACCG TGACGTTTTC CAAGCCTTTC TTGAAATTGC AGGTTATGGC | 5460 |
| GAAGAAGGGA TTCGTGTACG CTTTGTTACA GAAGATGAAG TTGGTGTCTT TGGAAAAATT | 5520 |
| GTTTCTTTGA TTGTAGAAGA AAATTTGAAT ATCTCCCATC CAGTCAATAT TCCGCGTAAG | 5580 |
| GATGGTAAGG TGATTATCGA AGTGCAAATC GATGGATCAA TTGATTTACC AGCCTTGAAA | 5640 |
| GAAAAATTTG AAGCAAATGG TATTCAAGTG GAAGAAATCG CTCGCACTTC AGCAAAAGTC | 5700 |
| TTGTAAGAAG GGAAGCCCAA AGGCTTCTTT TTTCATGAAA AGGGGATTAG AGCAAAAGAT | 5760 |
| GGAAAGAAAT GATAAAATAT GCTATAATGA AATAATGTAA AAAAGGAGTA TTTATGGACA | 5820 |
| TTTCAGTAAT TCGTCAGAAA ATTGACGCAA ATCGTGAAAA ATTAGCTTCT TTCAGGGGGT | 5880 |
| CTCTTTGACC TCGAAGGGCT AGAGGAAGAG ATTGCCATCT TGGAAAACAA GATGACAGAA | 5940 |
| CCTGATTTTT GGAACGATAA TATTGCGGCC CAAAAACGT CGCAAGAAAT AAATGAATTA | 6000 |
| AAAAACACTT ACAATACCTT CCATAAGATG GAAGAGTTGC AGGATGAAGT CGAAATTTTA | 6060 |
| TTGGATTTTT TGGCTGAAGA CGAGTCAGTG CATGATGAAC TGGTAGCGCA GTTAGCCGAA | 6120 |
| CTTGATAAGA TAATGACCAG CTACGAGATG ACTCTACTCT TGTCAGAACC TTATGACCAC | 6180 |
| AACAATGCCA TCTTGAAAT CCATCCAGGT TCTGGTGGTA CTGAGGCGCA GGACTGGGGT | 6240 |
| GATATGTTGC TTCGTATGTA TACTCGTTAT GGTAATGCTA AAGGCTTTAA AGTGGAAGTG | 6300 |
| TTGGATTACC AAGCAGGTGA TGAGGCTGGT ATTAAGTCGG TAACTTTATC ATTTGAAGGG | 6360 |
| CCTAATGCCT ATGGTCTCCT CAAGTCAGAA ATGGGTGTTT ACCGCTTAGT GCGAATCTCA | 6420 |
| CCATTTGACT CTGCCAAACG TCGCCATACC TCTTTCACAT CTGTAGAAGT GATGCCAGAA | 6480 |
| TTGGATGATA CTATTGAAGT GGAAATCCGT GAAGATGATA TCAAGATGGA TACCTTCCGT | 6540 |

750

| | |
|--|------|
| TCAGGTGGTG CCGGTGGACA AAACGTCAAT AAGGTTTCAA CAGGTGTACG TTTAACCAC | 6600 |
| ATTCCAAC TG GAATTGTTGT CCAATCAACA GTAGATCGTA CCCAGTATGG AAATAGAGAT | 6660 |
| CGTGCCATGA AGATGTTGCA GGCTAAGCTC TATCAAATGG AGCAAGATAA GAAGGCTGCG | 6720 |
| GAGGTAGATT CTCTCAAAGG TGAGAAAAAG GAGATCACTT GGGGAAGCCA AATCCGTTCT | 6780 |
| TATGTCTTCA CGCCTTATAC TATGGTAAAA GATCACCGAA CTAGCTTTGA GGTGCTCAG | 6840 |
| GTAGATAAGG TTATGGATGG GGACCTAGAT GGTMTTATCG ATGCTTATCT CAAGTGGCGA | 6900 |
| ATTAGCTAAG ATAGAAAGGA ACTCACATGT CAATTATGA AATGAGAGAT GTCGTTAAAA | 6960 |
| AATACGACAA CGGAACAAC GCTCTACGCG GTGTTTCGGT TAGCGTTCAA CCGGGGGAAT | 7020 |
| TTGCTTACAT CGTAGGACCT TCAGGAGCAG GGAAGTCAAC TTTTATTCGT TCTCTGTATC | 7080 |
| GTGAAGTAAA AATCGATAAA GGAAGCCTAT CAGTTGCTGG TTTTAATCTG GTTAAGATCA | 7140 |
| AAAAGAAAGA TGTCCCGCTT CTACGTCGTA GTGTTGGGGT TGTCTTCCAG GATTATAAAT | 7200 |
| TGTTACCAAA GAAACTGTC TATGAAAATA TTGCTTACGC TATGGAAGTA ATCGGGGAAA | 7260 |
| ATCGCCGTAA TATCAAAAGA CGAGTGATGG AAGTTTGGGA CTGTTGGA TTGAAGCATA | 7320 |
| AGGTTCGTTT TTTCCCAAAT GAACTCTCAG GTGGGGAGCA ACAGCGGATT GCGATTGCGC | 7380 |
| GTGCAATTGT AAATAATCCC AAAGTATTGA TAGCTGATGA GCCAACAGGA AATCTGGATC | 7440 |
| CGGATAATTC ATGGGAAATT ATGAATCTCT TGAACGGAT TAACyTACAA GGAACAATA | 7500 |
| TTTTGATGGC GACTCATAAT AGCCAGATTG TAAATACCTT GCGCCACCGT GTCATTGCCA | 7560 |
| TTGAAAATGG CCGTGTCGTT CGTGACGAAT CAAAAGGAGA GTATGGATAC GATGATTAGT | 7620 |
| AGATTTTTTC GCCATTTATT TGAAGCCTTA AAAAGTTGA AACGAAATGG TTGGATGACA | 7680 |
| GTAGCTGCTG TCAGTTCAGT CATGATTACT TTGACCTTGG TGGCAATATT TGCATCTGTT | 7740 |
| ATTTTCAATA CAGCGAACT AGCTACAGAT ATTGAAAATA ATGTCCGTGT AGTAGTTTAT | 7800 |
| ATCCGAAAGG ATGTGGAAGA TAATAGTCAG ACAATTGAAA AAGAAGGTCA AACTGTTACA | 7860 |
| AATAATGACT ACCACAAGGT ATATGATTCT TTGAAGAACA TGTCTACGGT TAAAAGTGTT | 7920 |
| ACCTTTTCAA GTAAAGAAGA ACAATATGAA AAATTAACCG AGATAATGGG AGATAACTGG | 7980 |
| AAAATCTTTG AAGGAGATGC CAATCCTCTC TATGATGCCT ATATTGTAGA GGCAAACT | 8040 |
| CCAAATGATG TAAAACTAT AGCCGAAGAT GCTAAAAAAA TTGAAGGTGT CTCTGAGGTT | 8100 |
| CAAGATGGCG GTGCCAATAC AGAAAGACTC TTCAAGTTAG CTTTATTTAT CCGTGTTTGG | 8160 |
| GGACTAGGGA TTGCTGCTTT GTTAATTTTT ATCGCAGTTT TCTTGATTTC AAATACCATT | 8220 |
| CGTATTACCA TTATTTCCCG CAGTCGCGAA ATTCAAATCA TGCCTTGGT CGGAGCTAAA | 8280 |
| AACAGTTATA TCCGTGGACC GTTCTTGTTA GAAGGAGCCT TTATCGGTTT ATTGGGAGCT | 8340 |

751

| | |
|---|------|
| ATCGCACCAT CTGTTTGGT CTTTATTGTT TATCAAATTG TTTACCAATC TGTCACAAA | 8400 |
| TCGTTGGTAG GGCAAAATCT ATCCATGATT AGTCCAGATT TATTTAGTCC GTTGATGATT | 8460 |
| GCCCTACTAT TTGTGATTGG GGTTTTCATT GGTTCATTGG GATCAGGAAT ATCCATGCGC | 8520 |
| CGATTCTTGA AGATTTAGGT AAAATAGCTG CTTTATGAG GAGATTGTAA AATCTCCTTT | 8580 |
| TTTGCTACAA GAGTTTTTGA AAAGAGATGC GCAGAAGAAA AGAGCTTCCA AAGAAGTCCC | 8640 |
| CCAGAGAAGA CTTC | 8654 |

(2) INFORMATION FOR SEQ ID NO: 99:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19718 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 99:

| | |
|--|------|
| TGTCGCGTCA AAATCATTAC TATGGCTATG TATAGCCCTT ACTATGACTT GGCTAAACAC | 60 |
| GTTTCGCTTTC AAATTTCTAG GCTCAGGCTG AAACAGTCTC CCAGGCTGTT CACTCCCGAA | 120 |
| TGCTAAAATC GTTCTTGATC GCTTTCACAT TGTACAACAT CTTAGCCGTG CTATGAGTCG | 180 |
| TGTGCATGTC CAAATCATGA ATCAGTTTCA TCGAAAATCC CATGAATACA AGGCTATCAA | 240 |
| GCGCTACTGG AAACTCATTC AACAGGATAG CCGTAACTG AGTGATAAGC GATTTTATCG | 300 |
| CCCTACTTTT CGCATGCACT TAACAAATAA AGAAATTCCT GACAAGATTT TAAGCTATTC | 360 |
| AGAAGACTTG AAACACCACT ATCAGATCTA TCAACTCTTA CTTTTCCTACT TTCAGAACAA | 420 |
| AGACCCTGAG AAATTTTTCG GACTCATTGA GGACAATCTG AAGCAGGTTC ATCCTCTTTT | 480 |
| TCAGACTGTC TTAAAACCT TTCTCAAAGA TAAAGAAAAG ATTATCAACG CCCTTCAACT | 540 |
| ACACTATTCT AATGCCAAAC TGGAAGCGAC CAATAATCTC ATCAAACCTA TCAAGCGCAA | 600 |
| TGCCTTTGGT TTTCGAAACT TTGAAAACCT CAAAAACGG ATTTTATCG CTTTGAACAT | 660 |
| CAAAAAGAA AGGACGAAAT TTGTCTTTC TCGAGCTTAG CTGACTTCAA CCCACTACAG | 720 |
| TTGACAAAGA GCCTAATTTC CATAAAATTT GACATGGAAA TTATAAAACC ATTACTAGTT | 780 |
| TAGTCCTTTT TGATAACGTG CCAATTCGGC TTGGTTCGCC CAAACATAGT GACCTGGACG | 840 |
| GATTTCTACC ATAGATGGCT TATCAGTCTC ATAGTCGTGT TGACTTGGAT CGTAAACCTT | 900 |
| CAAGACCTTC TTACGTTCCA AGATTGGATC TGGGATTGGT ACCGCTGAAA GCAAGGCTTG | 960 |
| AGTATATGGG TGAATTGGAT TGTAAACAA TTCTTCTGTT TCTGCAACCT CTACAATAAC | 1020 |

752

| | | | | | | |
|------------|-------------|-------------|-------------|------------|-------------|------|
| ACCCTTGTA | ATAACTGCGA | TACGATCTGA | AATAAAGCGA | ACAACCGACA | AGTCATGGGC | 1080 |
| GATGAAGAGA | TAGGTCAGGC | CGAGCTCTTT | TTGGAATTTT | TTGAGCAAGT | TCAAGACTTG | 1140 |
| GGCACGTACA | GAAACGTCCA | AGGCTGAAAT | TGGCTCATCT | GCAATAACAA | AGTCTGGTTG | 1200 |
| CATGACCAAG | GCACGGGCAA | TACCGATACG | TTGACGTTGA | CCGCCTGAGA | ATTTCATGAGG | 1260 |
| GTAACGAGTC | AAGTGCTCAG | CAAGAAGACC | TACTTCACGG | ATAATATTTT | GAACCTTCTC | 1320 |
| TTTACGTTC | TCTTCATCCT | TAAATAAACG | GTGATTGTAA | AGACCTTCAG | AAATAATATA | 1380 |
| ATCAACAGTC | GCACGTTTAT | TCAAACCTGC | GGCAGGGTCT | TGGAAATCA | TCTGGATTCTG | 1440 |
| ACGAATCAAT | TCCGCAGCTT | GTTTCACGCGA | TTTCTTACCA | TTAATCTTTT | GACCATCAAA | 1500 |
| AATGATATCT | CCATTACTTG | TATCATTTAG | ACCGATGATA | GCACGACCAA | TAGTTGTTT | 1560 |
| CCCACTACCG | GACTCACCTA | CAAGCGAGAA | AGTTTCTCCC | TTGTTGATAA | AGAAGTTAGC | 1620 |
| ATTTTAAACC | GCACAAACT | TCTTACTTCC | TTCAACCGAAG | GAAATTTCTA | AATCTTTGAT | 1680 |
| TTCTACTAAT | TTTTCAGACA | TTTCCTTCCT | CCTAGTCAGC | CAGATGGGCA | AATCCCATT | 1740 |
| TTTCACGGAT | CTTATCATGG | AGATTGCAA | TCACAGCTGG | TTTTTCTACT | TTCGGAGCAT | 1800 |
| CCTCATGAAG | AAGCCAAGTT | TTAGCCCAAT | GTGTCTCTGA | TACTGAGAAT | TGAGGAGCTT | 1860 |
| TTTGTTCGAA | GTCAATCTGC | ATTGCGTAGT | CAGAACGCAA | GGCAAAAGCA | TCCCCTTTCA | 1920 |
| GGTCAGTATA | AAGTGACGGA | GGTGTTCCTG | GGATTGAGTA | AAGATCCCCT | TTATCATCAG | 1980 |
| CAAGCTGAGG | CAAGCTAGAC | AAGAGACTCC | ATGTATATGG | ATGGCGAGGG | TCATAGAAGA | 2040 |
| CTTCCTCAAC | CGTTCCATAC | TCAACGATTT | CTCCTGCATA | CATAACCGCT | ACCTTATCCG | 2100 |
| CAATACTTGC | CACCACACCA | AGGTCGTGGG | TAATAAAGAT | TGTTGTGAAA | TGATACTCGT | 2160 |
| TTTGTAAAGA | TTTTAGCAAA | TCAATAATCT | GAGCTTGAAT | AGTTACATCC | AAGGCAGTTG | 2220 |
| TTGGCTCATC | ACAGATCAAG | ACATCAGGTC | GGCAGGCAAG | GGCAATAGCA | ATAACGATAC | 2280 |
| GTTGACGCAT | TCCTCCAGAA | TATTGGAATG | GGTATTTCATT | AAAACGTCTA | TCTGCGTCTG | 2340 |
| GAATGCCAAC | CTTATTTCATG | TAGTCAATGG | CCAATTCTTT | CGCTTCTTTA | GCTGTTTTC | 2400 |
| CTTGGTGT | TACAATAACT | TCTGTAATCT | GACTACCAAT | TGTTTAAATG | GGGTCCAAAC | 2460 |
| TAGTCATTGG | GTCTCGGAAG | ATAGTCGCAA | TCTTAGCACC | ACGAATTTGT | TCCCAATCCT | 2520 |
| TGTGAGAAGA | TAAAGCTGTC | AAGTCCTGAC | CACGGTAGTC | AATACTACCT | TGGGCAATAC | 2580 |
| GACCATTTTC | TTGAGCATA | CCTGTGAAGG | TCTTTGTCAA | AACAGATTTA | CCTGATCCTG | 2640 |
| ACTCACCTAC | CAAGGCTAAT | ACTTCTCCTT | CGACTAGTTC | AAGGGAAACG | CCGCGAATGG | 2700 |
| CTGTCAATAC | TTTGTACGGA | ACGTCAAAT | CCACGACAAT | ATCGCGAGCA | GTCAAAATTA | 2760 |
| CATTTTTTTC | TTTTGTTCATT | TCTACTCCTA | TCTATGTGTA | CGTGGATCAC | TAGCATCCGC | 2820 |

753

| | |
|--|------|
| TAAGTTTGA CCAACTACGA AAAGGGACAA GGATACCAAG ACAAGGGTTG TCAATGGAAT | 2880 |
| CCAGAACAAG TAAGCATTTGG TTGTTACGTT TTGTGAATAA TCCGAAATCA AACGACCCAA | 2940 |
| ACTTGGCACT GTAATCGGTA ATCCAAGACC GAAGAAAGAC AAGAAGGCTT CGTATGAGAT | 3000 |
| AAAGCTTGGG AGCATTGAG TCATGGTTGT CACAATAACA GATACCAATT GAGGCATGAT | 3060 |
| ATTTTGGCA ACAATCTTCA AGGTGGTGT TCCCAAAGTA CGTGACGCCA AGTTGTATTC | 3120 |
| CAAGTCACGA TAGCGCAAGA TTTGCACACG GATCATGAAG GCAATACCAA TCCATGTTGT | 3180 |
| TACGCTCATG GCAAAAATCA GATTCCAGAA TCCAGCTCCG ATTGAGTAAG TCAAGACAAT | 3240 |
| AACAATCAA AGAGGTGGG TGTGAGAT GACGTTGTAA ACTTCCATCA TGACACGGTC | 3300 |
| AACTGATTT GAAATACCCC AAATACCACC GACAAAACA CCGATAACCA AGTTAATCAC | 3360 |
| TGTCGCAATC ACAGAAATGA GGATGGAGTT ACGAGCTCCG AACCAGACAC CGTCAAAGAG | 3420 |
| CGATTTACCG TTACTGTCAG TACCGAACCA ATGCTCCGCA TTTGGCTTGA TATAACGAAC | 3480 |
| ACTAAAGTCG TTTACCTTGC TGACATCATT GAAATCAAAC TTAGAAAACA TTGGGTAGAT | 3540 |
| GAACTTATC AAAATGATGG CTACCAAGAT TCCCAACATG ACTACAGTTG ATTTTCTCTT | 3600 |
| CATAAATGT TTAACACTG ATTTCCAGTA AGAATATGCT GGCGCATCAA TAGTTTCAGA | 3660 |
| GGCAAATCG TCACGTTTA CAACTGAAA TTTTCTTTA TCGATTGTAG ACATTATTTG | 3720 |
| CCTCCTTCT CAGTCAATTT AATACGTGGG TCAATAATAG TCATCCAAAT ATCTCCCAA | 3780 |
| AGACGTGAGA AGATAGAAAT ACATGTAAAG ATGAAGACAA GACCAACGAC CATAGAGTTA | 3840 |
| TTAGATGCTT TTACAGAGTC AATCAACATT TTACCCATAC CTGGGAAGGC GAAGACTGTT | 3900 |
| TCAGTAAGGG TTGCACCACC GATAACCCCA ATAATGGCAG CAGGAATTCC TGAAACCAGC | 3960 |
| GGAACCATGG CATTTTAAA GATGTGTTG TTTGAAATT CTTTTTCAGA CAAACCTTTT | 4020 |
| GCACGAGCGA AACGAACAAA GTCTTGAGAT TGCAAGTCAA TCATGTAACG ACGAATCCAA | 4080 |
| ATGGCTGTAC CAGGAGCACC CAACAAACCA AGGATGACTG CTGGTAAAC GTAAGAACGC | 4140 |
| CAATCTCCAG CTCCAAGAT AGGGAATGAA TCTGGAAGG CAATAGATGA TCCAATCAAT | 4200 |
| CGAACGATGT AAACCAAGGC AATCGTTGGA AGAGCAAGCA AGAAGGTCAA AGCCCTGTT | 4260 |
| GAGAGGCTAT CAATCCAAGT GTTCTTGAAA CGAGCCATGG CTGAACCAAG TGGCACGGCA | 4320 |
| AGAGCATAGG CAAGAACCAA ACCAATCAA CCAGTAATAG CAGAGCTGAC AATCATAGAT | 4380 |
| GGATATTGGT AATTACTTTC AGTCGCTGTA TAAGGATCAT CTTTCCATA GCTAGCTACT | 4440 |
| TCACGAGAGT CAGCTGACT AGGTGACTTG TAGGTTCTTG AGTAAATATT TACAGAAGAC | 4500 |
| GTTTTCTTAC CTGTGGGAA CTGAACCTTG GCAGTTTGG TTTGTCCTTG ACCTGAGTA | 4560 |

754

| | |
|---|------|
| ATAACCTGAA GAACTGGTGT ATTAGCATAG GTTGGGTAAG AGTCACCTAA ATTCAAGTTC | 4620 |
| ACAAAGTTT GATGAACAAA TGGGAAGTGA CTGTTAAAGT ACAAGAGATA TTTATGTTTA | 4680 |
| GTTCCTGAAC CGACCAATGA CCATCCGATA GCTGGATCAT TTTCAAAACG AAGGTAGCGT | 4740 |
| TTCAAGTCTG GATTTTCAGG GTCTTGGATT TTATTGTAT GGTCAATGTC AATCAAGTTA | 4800 |
| GCATAGAAGT GAAAAACACG TTCAAAAATT GGAATTTTAC GACTAGCATA GAATTGACCA | 4860 |
| CTTTCAGTAA ATCTCTCCAA AGTCCAACCA TGACCTAATT GATTGATGTA CTTTTCATAA | 4920 |
| ATAGCTTTAT TGGTCGCATT TGCTTCTACT GTTACAGAAG AATCCATGCT ACTTGCCTTT | 4980 |
| TCTTGCAACT CTTTAGTATC GTAATACTCA ATGTAGCCCA TACGCTCAA CACAGTATTT | 5040 |
| TCATAGTTAT CACGTTTATC AGCCGTTGTC GCAATTTTAT TATAGTTAGG ATCTGCTTG | 5100 |
| AAAATCAATT TTCGAGGAAC CAAGGTATAG ATAATCGTGT AGGTCAAAGT CGTTACTAAG | 5160 |
| AAAATCGAAA CCAATGACCG CAAAACACGC ATAAAAATAT ATTTTTCAT ATTATTTCTT | 5220 |
| TTAAAAATCC CAAAAGAACC TTCTCCTCAT GGAGAGAAAG TTCTATTAGA AATTATTTAC | 5280 |
| TTACATGAC TTGCCAATC TTTTGTAGCT TTCTCATTG ATTCAGCTTT TTCTTTCAAC | 5340 |
| CATTTTTCAC GAGCTTTTTC ATACTCTTCC TTAGTACCA CTTTATCTTG TGATTTCAA | 5400 |
| TATTTGAAGT AAACATCTGA CCCCTTAGAG CCTGTTTGGC CAGAAGCTCC AGTAAATGGA | 5460 |
| ACAATTCGTG AAAGCACTGG TGCTGCACCA GAAGAAGCCA TAGCAGGAAT AAAGAGTGAA | 5520 |
| CTATCTGTCA ACCATGCTTG AGCCGCTGCA TATTTTTCAT AACGGACATT CAAGTCGCTT | 5580 |
| GTCTCTCTGG CAGCTTCATC AACTAATTTA TCGTATTCTT TCAAACCAAC TTGAACTACT | 5640 |
| GAAGGGCTAT TTGGATTATC AAATCCTAAA TATGTTTTTG TAGTTTCACT GCTAGTTGTT | 5700 |
| TTTAAATAT CCAGGTAAGT AGATGGGTCT TGATAGTCTG GCCCCATGA AACTCCTCCT | 5760 |
| GATACATCCC AATCCTCAGA TGAAGCATTG GCAGCATAGT AAGTAATATT AAGGAATTCA | 5820 |
| TCACTTGTCA TTTGTTGAAT ATCAACAACG ACATTTTCAA CACCAAGAAC TGTTCCTACA | 5880 |
| GATTGTTTAA AGGACTGAAT ACGAGATATG TAGTTTTTTG ATGCTTGGTC TACTGGAACG | 5940 |
| TCCAGATGAA TAGGAACTG AACGCCGTCT GCTTCTAAAG CTTTCTTAGC TTTCGCAAAAC | 6000 |
| TCTGCCTTGG CCTGTGCAGC ATTGAATAAA CCATCCTGCC CATCAGCTAA ATTCACACCT | 6060 |
| TTCCACTCAT CACCATAAGC AGGAAGTTGA GCAGCGACTA AATCACCAAA GGTCTTCTCA | 6120 |
| CCAGCTGAAA CAAAGTCTGG TTTTACAAAT AAATTACGAA CTGCTAAAGC TGCTCCATCT | 6180 |
| TTACCATTGA TTTGAGCTGA GTAAGCTGAG CGATCAAGAG CAAAATTCAA GGCTTGACGG | 6240 |
| AAATCTTTGT TAAGCAATGC CTTCTTAGTA GCTACTTTCT CTGAATCTGT AGTTTTAGAA | 6300 |
| GTATAGTTGT AACTTTGGCG ATCAATATTC ACACCCAGAC CAGCAATCCC AGAGCCTGAT | 6360 |

756

| | |
|--|------|
| CGTTTCCCCG ACTCCTGACT CATATCCATC ATCAAGCGAA CAGGAGCAAC AGAAGACAAA | 8160 |
| ACTAATAAAA TAGTCCCCAC AATTCCGTAA CTCAGAATCG TATCAATATA AAGACTGTGG | 8220 |
| GCATGTTTCAT GATAAGGAGC ATGTATCCGA GGATAAGAGT TCATATAGGT CAATGGCCCT | 8280 |
| TCACCCCAAA AAGGATTTTG CTAAACAAG GCCATCCCAG CATCCCAGAT AGAAATGCGT | 8340 |
| TCTTCCATAG AAGAGTCTAA AGTACCCATT CGAACTCCCA AATCACTAGA AAAGAGGAAA | 8400 |
| CTCAAACCAA TCGCGAAGAC CCCAATACTA AGCCAAAAGG CCTTCCAGTT TTTAATAGTC | 8460 |
| GTAAGAGAT AGATAATTGC TCCAGCGATA ATAGCAGGAA AGGCAGTTCG ATTTTGAGTA | 8520 |
| AAGTTCAAAC CAAAGAGATT AACAAAGCCT GCAATCACAC AGAATACTTT CAACCAATTC | 8580 |
| AACPTGGTCG TTGTAACAG ATAGAAAGCA ATCATAATAC AGAAACAACA AATAATTCCA | 8640 |
| TAATAATTAG GATTAAAGAA GGTCACTTCT GCCCGTTCT GATGCCACAC CTGCATATTG | 8700 |
| GGTGAAAGAA AAGCATAGTT AAATTCTTTC ACAATTGGA AATGTTCTAA ACTGGCAAAA | 8760 |
| GCAGCTGACA AGACACTACC AAACAAGACA AACTGCAAAA TCAATCGAAA GAATTTATGG | 8820 |
| GATAAATCG ACTGATAGTG CAAAAGAAA ATAGTAAATA GAAACATTCC TACTGAAGCC | 8880 |
| ACAAGACCCA TCCAATTTTG TGCAAGAATG GATATAACAG TACTATAGCT AAGAAAAAGA | 8940 |
| AGCAGCATCG GATGCTCCCC CATTTTCTGA AGAATACTTT TCATGTCTCC TGTAAAAATC | 9000 |
| AAACTGATAA TATATAACA GAGTACAACT AAAAAAGAT AAAAGGGTAA AAAGATACTC | 9060 |
| AGGATAATTC CCAATAAAAT CAGCTCTTTC CTAGACAACC CCTTCAGCTT TTCAATAAAG | 9120 |
| CCTATTGATT TCAAAATGAA TCCTTTCTCT CCAAAATCAGC TGATTCAGAT AATAGTAAGC | 9180 |
| TATCCTATAT TGTACCACTT TTTTAGCAAT TTGAAAACAA AGGAAACGTT TTCCAAAATA | 9240 |
| AAAACCTAT TTTATCCACC ATATCAAGGC TTCAAAATGA TACTTCAACT CCATTCTCAA | 9300 |
| TTACCCGATA AGTCTGATTT TGCAAAATCA TTTCTACTAC TGCTGTTACG GACTTATCTT | 9360 |
| TATTTTGACG TTTGATTACA ATGCTGTGAG CTGTTGGTGT CTCTATCTCA GTAGTCCCTT | 9420 |
| CTAGATCAAA GGCTTCTGAA CGGTTACGGA AAGAAAATAG ATTGAGAAGG GCCTTCACAA | 9480 |
| CAGGTCGTTG CACTTCTTTT GCTATTTTCT CGTTGCTATA GTAATGACGA TTAATATTTT | 9540 |
| GACCTTCTTT AGTTTCTTCT AATAATTTCA AGTCATTCTT GCCTGCTAAT AGACCCACAT | 9600 |
| AGTAAATCTG AGGAATACCT GGGGCAAAAG CTTGAATTAG ACGAGCGAGA AAATACTTGA | 9660 |
| CATCATCATC TCCAAGCGCT GAATAGTAGG TTGAATTGAT TTGGTAGATA TCTAAGTTGT | 9720 |
| TATACTCGGC ACTAGAGTAC TTACGTTTGA CATTGGCTCC AACCTTATAG AGTTCATTTG | 9780 |
| AAGCATAGTC AATCTCCTCA TCGGTCAGGA TATCCTTGAC ATCTACTACT CCAATCCCAT | 9840 |
| CATGGGTATC TAGCGTCGTA AATTGCTTCA TCGGGCTCAT CTTTAACCAC TTAGCCAAAC | 9900 |

755

| | |
|---|------|
| TGTGTGTAAT AGATATTGTC CTTGTATTCT TCTGCAACCT TAGAATAGTT GGAGCTGGTA | 6420 |
| GGGTAAAGAC GGGCATAACT ATAAGCTCCA CTAGTGAAGT TACGCTCTAG CGACTCCTGA | 6480 |
| TCTGATCCAT CATAGTAAGC TAGATTGATA GTATCTAGGT GGACATTTTC TTTATCCCAA | 6540 |
| TATTGCTCAT TTTTACAAA CTCTACAGAA GATTTTGCAG TCAACCCTTT CAACAAGAAT | 6600 |
| GGACCATTAT AAAGCAAGGA TGTCGGATCT GTTGGTTAG CAAAATCGCT TCCTTTTGAT | 6660 |
| GTTTCGAATT CTTCATTCAG AGGCCAGAAA ATAGAATAGG TCAACTTAGA GTTCCAGAAC | 6720 |
| GGTTCAGGCT GGTTCAAAGT GTATTGTAAC GTATAATCAT CAACCGCCTT GACACCAACT | 6780 |
| GTTGAAAAAT CTGTTGAAGT TCCTGATAGA TAATCTGCCA AGCCTTTAAC CGAATTTTCA | 6840 |
| GCTAAATACA TAGCTTCTGA TTTTATCTCT GCTGCGTGTT TTAAACCGTT CACGAAATCT | 6900 |
| TTAGCCGTCA CCTCTGCATA TTCTTCTCCA TCAGAGGTAA ACCATTTAAC CCCTTTACGA | 6960 |
| ATCTTATAAG TGTAGGTCAA ACCATCCTTA GAGACTTCCC AATCCTCTGC AACTGCAGGA | 7020 |
| GCAAGATTAC CGTAATTATC GTTAGTGAAT AAACCATCAA TCCCATTTGA AGTCACTACT | 7080 |
| GTTGTACTAT TTTTACTTGA AATCAGGTAG TCCAAGGTTT CTGGGTCTGC TGTATAAACA | 7140 |
| TAGCCATAAG CTTTAGGGGC TGATGAATCA GATGATTTTG AAGAACTGCA TGCTGCAAGT | 7200 |
| ACACCTGCTG CTAATAAAAC AAGACCTGCT GTAGCAAATA CACGATTTT TTTTATTTC | 7260 |
| TACTCCTCTG TTTATGTGAA TTATAGATTG ACAACCATTA TATCACATTA TCCATTAAAA | 7320 |
| ATCAAAACAA TTTTCAGAAT ATTTAGGCTT GTTGGCACAA ATTTTTCATT TTTTGTGAAT | 7380 |
| ATATGATTCA AATTGTCGTT CGAAGTGTC AAGACTACAG TGAAAATAGG AAATTTGACG | 7440 |
| CAGAAACTTT GGAGTTTAGG AAGACATACA GTAAAATGAA ATACGGACGG AACAATGTGA | 7500 |
| TTTGGGAAT CAAATTAAAT TATAACAATA TTGTAGAAGT ATCATTCTAG TATTCAAGAT | 7560 |
| TCAGTTTACT ATGTCTTTTC ACACCAACCT TATCCCGAAT TCAATTACTT TTGTGATTTA | 7620 |
| CATATATAGA TTAAGACTAT CTTTATACT TTAAAATTT TCGCTACCTT ATCCACTATA | 7680 |
| TGCTCCTCGC TATCACGTTT CTATTCATAG CCTACGATTT CACTATTGCT TTCTCTGACA | 7740 |
| ATTCTTATTT CTGCGTCAG ACTTAAAACG ATCTATCCCC AGACCATTTT AATCCGCTAC | 7800 |
| CTCACGATAG TCAGGCTTGG GGAGCGCTAT TGTATTCAAC GGTAGTGGAG CCCTACAGAG | 7860 |
| GACTTACACC TCAGATGCAC GACATGCCCA TCGTATAAAA AATCTCCTAC CCAAGGTAGA | 7920 |
| AGATTTCAAA CTTATAAAAC TTAATCCGTC ATGTCCGATA CCAACATTCG ATGCTCCAAT | 7980 |
| GGAATACTGC ACATAACTAG CAAGAAAATA AAGCTGACT GAATCCAGAA GAGAGCCAAG | 8040 |
| TCAAAAATTC CGTGCACAGC AACCCTGTA AGGAAAGATA GATAAAGGCC GATAATCGGA | 8100 |

757

| | | | | | | |
|------------|------------|-------------|------------|-------------|------------|-------|
| GCTCTGTTCT | GGAAGCTGTA | AGAGTATAAA | GTGTCACCAT | TGGAAGAGCA | AAATCATAAA | 9960 |
| CATAGTAATC | ATGGTCTGCT | ATTTTAAACT | GAATCGAATA | GTGTTTCATGA | ATCTCAGGTA | 10020 |
| AAAGCTCTGT | CCCATACTCA | GCAGCGATAT | CTCGAACTTT | GTCCAATAAA | TCCCAAATAT | 10080 |
| CTGGTTCCAC | AAAGAAATCA | TTAGTATCCA | ATTTCTTCAC | TGCATAAGCA | AAGGCATCTA | 10140 |
| GACGAATCAA | ATCACACCCA | TTACTTGCCA | AGTGCTGAAT | GGTCTTACGG | ATAAATTCCA | 10200 |
| TAGTTACTTC | TTGGTCCACA | TCAAGATCAA | TCTGCTCCTC | ACCAAAGGTA | TCCACAAAT | 10260 |
| GTCCACTGA | ACCATCTTCA | AACACAATCT | CTTGCTTTGG | TGCACGATCC | TTACGCTTGT | 10320 |
| AAATTAAATC | TACATCAGAC | TGTGTCGGAC | GGTTTTCTGG | CCAAAACCTA | TCCCAGTTTA | 10380 |
| AAAAGAGAGC | TTTAAATTCA | CTGGCTTCAT | GTTTTTCTTG | ATAGTCCTTA | TAATACTTGG | 10440 |
| ATTGACGAGA | AATATGATTA | ATCATAAAAT | CAAACATAAG | ATAATATTTT | TCACCTAAAC | 10500 |
| GCTTCACATC | CTCCCAATCA | CCAAAAGCTG | AGTCCACTTC | GTCGTAGTCA | ACTGGCGCAA | 10560 |
| ATCCACGATC | AACTGTTGAT | GGGAAAAATG | GTAAGAGGTG | AACTCCTCCA | ATAGCATCTC | 10620 |
| CAAAATGCTC | TTCCAAATTA | TCATATAAGT | CTTTAAGATT | ATTTCCAAGG | CTATCAGAAT | 10680 |
| AGGTAATCAA | CATGGTTTTA | TTTTGAATTG | GCATCATTAC | TCTCCTTTTT | CTAATTGAAG | 10740 |
| CCAAGTCTCA | TATGATCTGG | CTTCATAAAT | AAAATTCATT | TTAAATCTCT | ATTATCATC | 10800 |
| AAACTCGTAC | TAATATAGAC | TGTGATAAAC | AAAGTACTAC | TTTCTTGTTT | TCTGCATAGA | 10860 |
| ATTATCAACA | AGCTAAACTC | TTCTCTGTG | TCAAAGACTA | TAGATTCAT | GAGCTCTTCT | 10920 |
| TATACTCTTC | GAAAATCTCT | TCAAACCACG | TCAGCTTCAC | CTTGCCGTAG | GTATGGTTAC | 10980 |
| TGACTTCGTC | AGTTTCATCC | ACAACCTCAA | AACAGTGTTT | TGAGCAACCT | GCGGCTAGCT | 11040 |
| TCCTAGTTTG | CTCTTTGATT | TTCAATTGAGT | ATTACTTCAC | TGCCCCGTG | CTCATTCCTG | 11100 |
| AAATGATATG | GCGTTGGAAG | AAGAGATAGA | CAATGGTGAT | ACTGATAATG | CCGACCACGT | 11160 |
| AAGAGGCAAA | GCTTGGTCCG | TAGTCGTTGA | AATATTGGCC | TGCGTAGTTG | TATTGGAACA | 11220 |
| AAGGCAGAGT | CCACATTTTG | GAATCCCGGT | TCAAGACAAG | GAGTGGCAAC | ATGAAGTCAT | 11280 |
| TCCAGAACCA | AAGGGCATTG | ATGATCATGG | TTGTCGCATG | CATCGGTTTC | ATCATTGGGA | 11340 |
| AGATGATGCG | GAAATAGGTT | GTAATTTGAT | TAGCCCCATC | GATCTCTGCT | GCTTCATCCA | 11400 |
| GACTTTCTGG | AATCGAGATT | TTGATATAGC | CAACATAGAG | AAAGAGGGTC | TGTGGAATCG | 11460 |
| CATAGGTCAA | GTAGAGCAAG | ATCAAACCAA | AGGTATTAGC | CAAACCGAGT | TTACTCATCA | 11520 |
| TAACCGTAAT | CGGAATCATG | ATGACTTGGA | AAGGTACGAA | GATTCCGAGG | ATTAAGAGGG | 11580 |
| TATACATGAT | GGTAAAGGCT | TTTCTTTTAC | TCATATTGCG | AGCGATGGAG | TAGGCTGCCA | 11640 |

758

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-------|
| TAGGGATAAA | GATCATTACT | GCAAGTAAAG | ACAAGACAGT | GATGACGACA | GAGTTCCAAT | 11700 |
| AATAGCCTCC | AATCCCATCA | GCTAAGAGAC | GGCTAAAGTT | GTCCCATGTG | AAGTTGGTTG | 11760 |
| GAAAGCCAAA | GAAATTATCT | ACAATATCCT | TAGTGGGTTT | GAAGGAACTA | AAGAGGGTAG | 11820 |
| CAAGGAGCGG | CACTAAAATC | AGAACCGATC | CTAGAATCAA | TAGAATGTAT | TTGCCAATCA | 11880 |
| GGGCTTTTCT | TTCATCTTGT | TTCATCATGC | TTCTCCTCTT | AAATTTCAAA | TTTCTTAGAT | 11940 |
| ACTCTCAATT | GGATGATCGA | AATCACTACA | ATTAAGAAGA | ACAAGATTAC | GGCAATGGCA | 12000 |
| TTGGCATAAC | CGAATTGGTT | GTTTTAAAG | GCATAGTTAT | AAACCAAGAG | CCCAAGTGAG | 12060 |
| GTGTGGCAT | TGTTGGACC | ACCACCGGTC | ATGGCAAAGA | CTTGGTCAAA | GGCAGTCAGC | 12120 |
| CCACCTTTTA | GGGCTAGGAT | AAAGACCATA | GAGACACTTG | GTAGCAAGTA | AGGCAATTCA | 12180 |
| ATGTTCCAGA | AACTTGCTT | GCTAGTCGCA | CCATCAATCC | TTGCTGCCTC | TGTAATCTCA | 12240 |
| GTMGGAATAG | ATTGCAAACC | AGCTAGGAAG | ATGATGATGG | GCATAGCCAC | CCCTTGCCAA | 12300 |
| AGAAGGACAA | AGACAGCCGC | AAAGATTGCT | CCCCACTTAG | TCCCTAAAAG | ACTGGTTTGG | 12360 |
| AAAAATCAA | TATGAAGGGC | ATTTCCAATC | GCTGGAAGAC | CGTAGTTGAA | GACTTGCTTG | 12420 |
| AAGATCAAAG | CCACTGTCAA | ACCAGATAAA | ACAGCTGGGA | AGAAGAACCA | AGCACGGAAG | 12480 |
| AAGGTTTGGC | CTTTGATTTT | AGAATTCAAG | ACACGCGCAA | TGAAGATCCC | GAGTGCAATC | 12540 |
| TCACCAACCA | CCATGGCAAT | CGCAATGATT | GCGGTAAAGC | CAATCGCATT | CATGAATTTT | 12600 |
| GGATCCATGA | AGAGGAGCTT | AAAGTTGTTT | AAGCCAACAA | ATTTGTAGTT | ATAAGTCAAT | 12660 |
| CCTGTCCAGT | TGGTAAACT | GTAAGGCT | CCTTGAAACA | TCGGCACATA | GAAGAAAATT | 12720 |
| GCTTGTAACA | AGAGGGGGAT | GACCACAAAA | GCCCATGCCC | AATATTTTTG | TAATACTTTT | 12780 |
| TTCATAGTCT | CTCTACTCCT | AATCCACATC | CGCTTTCATC | GGGTAAAGA | AGGCATTCAA | 12840 |
| ATCATTGACC | ATGCCTTGTT | TATCACCGGT | CAAGACATAG | TTCATGGTCA | AGGTATGGAA | 12900 |
| GTCTGCTTCA | CTGGTCCAGT | ATTGTTGCAA | CCAGACCAAG | TGACGATCCG | TAAAGGCATA | 12960 |
| TTCGGTCATA | CCAGCAAGCG | GTGAATCTTC | TCCTGCTTGT | TTGACCCCTT | CGATCGCTGT | 13020 |
| TGGAGATCCG | TCCACATCGT | AGTATTTTTG | CATGACTTCT | GGACGGGTCA | TATATTCAC | 13080 |
| AAAGGCATTG | GCTTCTTTTG | GATGTTTGGT | GGTGGCTGAG | ATAGACCATG | CCAAGTCTCC | 13140 |
| CGCACCAACG | GTTAAGCTTT | GTCCTTTTTC | TTTTCTGGA | ATCATGAAGG | TCCCAATCTT | 13200 |
| AAAGTTCGGT | TTTTGTTCAT | TAATCGCTGT | GATCGCCCAA | GACCCATTTG | GTGTCATGAG | 13260 |
| GACATCCCCA | CGTGCGAAGG | CTCCGATAAC | ATCGGTATAG | CCAGCACCTT | CCCAGTTCTT | 13320 |
| TTGCTTAGAT | CCATTGATGC | GAAGGATGTC | CATGACCTTG | ATATCATCTT | TCATAATCGG | 13380 |
| ATCCGACAA | TTAATGGCAT | TTGGTTGAGA | ATAACGAAGG | TATTGATTTG | CTTCTTTTCC | 13440 |

759

| | | | | | | |
|-------------|-------------|------------|-------------|-------------|------------|-------|
| TCCACCTGTT | GCTGTCGCAA | AGGCTAATG | ATTGTAACCA | TTGAGTGTCC | AAGCATCTGC | 13500 |
| ACCTGCAATT | CCAAATGGTG | TTTGTCTTT | AGCAACGATA | TCTTTGACTA | ACTGTTCAAA | 13560 |
| TTTCATCCCAG | GTTTCAGGAA | CCTTCAAGCC | CAGTTCTTCG | AATTTATCTT | TGTTGTAGTA | 13620 |
| AATTCCATAA | GCATTAGCTG | TAAAAGGAAC | GTTGTAAACT | TTTTCGTTTA | CAGCATATTT | 13680 |
| TTCAGCGTAG | CCATTTTTC | CGCGTTTCAG | GTAAGTCTTG | TTGCTCAAAT | CTTCAAAAAC | 13740 |
| ACCTGCTTTT | GCCCATTTCTT | GCAGTTCGAT | GGACTGTGGG | TAAATATGTA | CCACATCAGG | 13800 |
| CACATCTCCT | GCGAGAACGC | GTGTCTTCAA | TACTTCACCA | GCATTGGTA | CATTGACGAC | 13860 |
| TTTGACCTTG | ATCTTAGGGT | TTTCCTTCTC | AAAATCACGA | GTGATTTCTT | CCAAGGTTTT | 13920 |
| GGTCATTTCT | TTTTTCTGGT | TGAAATACTC | GATGGTCACT | GTGCCATCCG | CAGATTTACC | 13980 |
| ATAGTTGGAG | CAAGCGCCGA | GCCCAAACAA | AGCTAAACCT | GTAAGTTCGA | GAAGTCCGAT | 14040 |
| TTTTTTATAC | CATTCCATTA | GAAAGCCTCC | TTTATAAAAT | TATACACCCT | TATTGAACTG | 14100 |
| CACCCCAAAA | GTTAGACAGA | ATAAATCTAA | CTTTTGGGGT | CAGTACATAT | CATAGTTTTT | 14160 |
| TAAAAATATA | CTGTCTACTC | AAAAAATCTC | CTTGGGATAA | GATAACAGTT | AAGCCCGCAT | 14220 |
| ACATTAGTTC | TGCACCTGAG | TAAACTTCGC | CATTTTCCTG | TAATTTATAT | AGTCCCTCTT | 14280 |
| CATCCAAATC | TTTTAATTTT | AAAGTTGTTT | CCATGGTCTC | TACAACAGAT | AAAACGCGAA | 14340 |
| CGTAGGTAC | AATCGTTTGA | TTTCCGTAAT | TAAATTGTAC | AGCTGCTTCA | TTGGATACAG | 14400 |
| TATCAGGATT | AATTAGTCTA | TACTGCTGTC | CTAACTGAAC | TACTGGTCGT | AATTCTTTAT | 14460 |
| ACAAAGTCAC | CTGATTAGCA | ATCGTAGCTT | TCTCTTCATC | TGATAAAATTT | GTCAAATCAA | 14520 |
| GTTCATAGCC | CAAATTTCCC | ATCATTGCTA | CAAGGCCACG | TGTTTCTAAT | GGTGTCAATC | 14580 |
| GTCCCATCTG | ATGATTCGGT | ACTGCTGACA | CATGAGCCCC | CATAGAAATG | GTTGGATAGA | 14640 |
| GATAGGATGA | ACCGTATTGA | ATTGGTAAAC | GTGCAATGGC | ATCAGTATTA | TCACTAGCCC | 14700 |
| AGACTTGTGG | GAAATAGCGC | ATCATACCAA | GATCATTTTCG | TCCACCACCA | CCAGAGCAGG | 14760 |
| ACTCAAAGAG | AATATGGCTG | TGCTTCTCTG | TCAGATAAGA | AACGAGTTCA | TAAAGCCCCA | 14820 |
| GCATGFACTG | ATGAGATTGC | ATCTGTGTCT | CTAGATAAGT | TAATCCATTC | CCTAGCTTAG | 14880 |
| TGATATTGCG | GTTCATATCC | CATTTAATGT | AATCAATATC | ATGATAAAAT | AGGAGTTGAT | 14940 |
| CTAAGACACT | TTTCAAGTAT | TCTACTACCT | GAGGATTGGC | AAGATTAAGT | ACTAATTGAT | 15000 |
| TCCGAGAATA | AGTATGCTCA | TAGCCAGGAA | CCTGAATAGC | CCAGTCAGGA | TGTTGACGAT | 15060 |
| ACAAATCACT | ATCTACAGAA | ATCATTTCGG | GTTCTAACCA | AAGTCCAAAC | TGCAAACCTC | 15120 |
| TTTCATGGAT | AGCTGAAATC | AGACTTTCTA | GACTTCCACC | CAGTTTTTCC | TCATTAACAA | 15180 |

760

| | |
|--|-------|
| CCCAATCACC TAAAGCACGA TTATCATCAA AACGATTGCC AAACCAACCA TCATCTAATA | 15240 |
| CAAAAAGTTC AATGCCAACT TTCTTAGCTT CATCTGCTAA CTCTAACAGT TTTTCTCTCT | 15300 |
| GAAAGTCAAA GTAAGTAGCT TCCAGTTAT TGATTAGAAT TGGACGTTCT TTTTtagaaa | 15360 |
| ATTCACTTAG CATAATGTGC TTCAGTACAA AATTCTGACT TTCATGACTA ATACCAGTTA | 15420 |
| ATCCCTGATC TGAATGAGTC ACTAAAGCTA CCGGTGTTTC AAAGTATTCC TCAGGAGCTA | 15480 |
| ACTTCCAAGA AAAGTTTCTT GGATTAATGC CAATAGCCAC CCGAACTTCA TTCAATTGAT | 15540 |
| TTTTTTGAAC AAAAGCTTCA AAGTTGCCAC TATACATTAG TTGAATAGCA AACACATTCC | 15600 |
| CAGCATCCTC TGTGACTCCT TGTTCGCATA GTAGAAGAGC TGGTGTTTGA GCATGACCAG | 15660 |
| AAGCACCTCG GTTTGAACTA ATCGAAAAGA TTCCTTGTTT TACCTGTGTA CGTCTAACAG | 15720 |
| TCTTTTCACG AGCATAAGCA CCCTGCAGAG TTACTATTTC GTAATCTGCA GCTGGAAAAT | 15780 |
| CAGCCATAAA AGAAAAATCT TTATGGATGA CAACTTCCTG ATTACTATTA TTATCTAATT | 15840 |
| TACTGTAGCT AGCAATAGTC GCATCATTAT TAAAAGTAGT ATAATACAAA GTCAGACTAA | 15900 |
| GTTGAGCCTT AGAATCTTCT AACATTAAGA CAAGAGTCTC TGTATCGTCC ATGCTATGTG | 15960 |
| GAGAAGGTAA GCCCTGTGGA CCATTCTGAC CTTTAAAAAT CTTTGCTTCT ACAAAATCGAA | 16020 |
| AGTCTGTTAC TTCAGTTACA CTATGCTGAA CCTGTATGGT TGGTTTCCTA AAATCTCCTA | 16080 |
| AGCCATGTTG TCCAAAAATC TGTCGCTGAG TATCTAAACT AAAGGTCGA TTAGTAGCCG | 16140 |
| TTGGATTTCC TGAAAAGGCA TGGTCTCGTT CATAAACACT ATTGGAACCT TTATAGTTCT | 16200 |
| TAATAGTCTT TCCTAAATGT TTCAAAAGTA AGTAGCCATT TCGATTTTCA ATAATCAAAC | 16260 |
| TTAGATTTTT ACTCTCAACA TAAAATAGAT TATTCTCTAT CCTAACTCCC ATTTACTTCA | 16320 |
| CCTCATCACT TTATTGATTA TATTTTATCA CCTGAAATCG CTTTCCAAAA TAGAAAAATG | 16380 |
| TCTCAAGAAT ATGGTAAAAT GTTAGGTAGG AGGTAGCACA TGTTAGTTTT TTCAGAATAC | 16440 |
| CAGACTGGAA CAATCGACCT TGCCCTAAGC TTTTATGGAT ATGAGGAATG CACACCTAAT | 16500 |
| TACTCTTTTG GTCCAGCCAT TCGTGATACA TACGTTCTAC ATTACATTAC TAAAGGACAA | 16560 |
| GGAAAATTTT ATTACAAGGG TAAAATTGTT GATTTAAAAG AAGGAGATT TTTTCTATTA | 16620 |
| AAACCAGAGG AACTAACCTT TTATCAAGCA GATAGTAAAG AACCTTGGGC CTACTACTGG | 16680 |
| TTAGGAATCA CTGGAGGGAA AGCCCCTGAT TATTTTGCTC TTTCCCAAAT TTCTGATCAA | 16740 |
| TCCTATCTCA TCCAATCTGA AACTTGTCAT ACCCAGACTA CTGCAAACT CATCTCAGAC | 16800 |
| ATTGTCCGCT TCGCTCAGAT TACAAAATCA AGTGAATTAG CTCAACTCCA TATCATGGGA | 16860 |
| CAACTTCATG AACTGATGTT TCATCTGGGA ACTATTGCTC CCAATCAGAA AAAAAAGAAT | 16920 |
| ATTTTCATCAA CCCACCAACT CTATCTTGAA TGCAAACGAT TAATTGATAG CCACTATCCT | 16980 |

761

| | |
|--|-------|
| CAATCACTTA CAATTCAAGA TTTAGCAAAA GAACTATCCG TTCACAGAAG CTACTTATCA | 17040 |
| AGCGTATTCA AAGAATTAA TACCTTATCA CCCAAAGAAT ACCTACTCTA CGTTCGAATG | 17100 |
| CACCGAGCTA GACAACCTCT CGAAAATACC CAAGAGTCCA TCAAGGTAAT TGCATACTCG | 17160 |
| GTAGGTTTTT CAGATCCACT CCATTTTTTCG AAAGCTTATA AACAACTCTT TAATCAGACT | 17220 |
| CCAAGTCATA CAAGAAAAGA ATACTCTCAA TACCAACTAG TAAGAAAGGC AACATTATGA | 17280 |
| AATCCTACCA AGCTGTCTAC CAAATCCTAT CTAAGAAAC CGACTATATC AGCGGAGAAA | 17340 |
| AAATCGCAGA AAAACTATCC CTAAGCCGAA CAGCAATTTG GAAAGCCATC AAGCGACTAG | 17400 |
| AACAAGAAGG CATTGAAATT GATAGTATCA AAAATAGAGG ATATAAACTG ATGAATGGTG | 17460 |
| ACCTTATTCT TCCAGAGATT CTAGAAGAAA ATCTTCCAAT TAAAGTCAGC TTTAAACCCG | 17520 |
| AAACAAAATC AACACAATA GATGCAAAAG AAGCAATTGA TTTAGGCCAT GAAGCAAATA | 17580 |
| CCCTCTATCT AGCTTCCTAT CAAACAGCAG GCCGAGGCCG TTTTCAACGT TCCTTCTACT | 17640 |
| CACCACAAGG TGGTATTTAT ATGACACTCC ATCTTAAACC AAATCTCCCC TATGACAAAT | 17700 |
| TACCATCCTA CACACTACTT GTAGCTGGAG CTGTCTACAA AGCCATTAAG AACCTAACTT | 17760 |
| TAATAGATGT CGACATAAAA TGGGTCAATG ATATCTATCT AAACAATCAT AAAATGGAG | 17820 |
| GAATCCTTAC TGAAGCAATG ACCTCTGTAG AAAGTGGCTT AGTCACAGAT ATCATTTATTG | 17880 |
| GAGTAGGTAT CAATTTCACT ATTAAAGACT TCCCTCAGGA ATTAAAAGAA AAAGCTGCCA | 17940 |
| GCTTATTTAA AGCTACAGCT CCTATAACAA GGAATGAATT GATCATAGAA ATCTGGCGTG | 18000 |
| CTTTCTTCGA AACACCAGCA GAAGAGCTAT TATACCTATA CAAAAACAG TCATTCAATC | 18060 |
| TAGGAAAAGA AGTCACTTTC AACTAGAGC AAAAGACTA CAAGGGACTT GCTAAAGACA | 18120 |
| TCTCAGAAAA TGGAAAACCT TTAGTTCAAT GTGATAACGG AAAAGAAATC TGGCTAAATA | 18180 |
| GTGGCGAAT TTCTCTCAAT AGTTGGAAGT AAAATAACAC AATTATAATA TAAACGATAT | 18240 |
| AAAAATAACT TCAGATTAGT AATTCAATTA AGTTTACGG ATCTGAAGTT TTATTGGCTC | 18300 |
| TAAAAATAAA AAAGAGAGTT ACAGACTCTC ATTAAAACGG AGAATAAGGG ATTCGAACCC | 18360 |
| TTGGCCAGT TACCCGACCT AACGATTTAG CAAACCGTCC TCTTCAGCCT CTTGAGTAAT | 18420 |
| TCTCCAATTA ATGGGCACGA GTGGACTCGA ACCACCGACC TCACGCTTAT CAGGCGTGCG | 18480 |
| CTCTAACCAC CTGAGCTACG CGCCCAAGTT AAAAACTTG GTAATTTGAA CAAAGTTCAA | 18540 |
| AGCGGGTGAC GAGAATCGAA CTCGCGACAA CAGCTTGAA GGCTGTAGTT TTACCACTAA | 18600 |
| ACTACACCCG CATAAATACT ATCAATAAAA TGGCGCGAGA CGGAATCGAA CCGCCGACAC | 18660 |
| ATGGAGCTTC AATCCATTGC TCTACCAACT GAGCTACCGA GCCTTATTGC GGGAGCAGGA | 18720 |

762

| | |
|--|-------|
| TTTGAACCTA CGACCTTCGG GTTATGAGCC CGACGAGCTA CCGAGCTGCT CCATCCCGCG | 18780 |
| TTAATAATAT AAAAGGAGGA TGTGGGATTC GAACCCACGC ACGCTTTTAC ACGCCTGACG | 18840 |
| GTTTTC AAGA CCGTTCCCTT CAGCCGGACT TGGGTAATCC TCCAATATTC AAATGGACCT | 18900 |
| TGTAGGACTT GAACCTACGA CCACTCGGTT ATGAGCCGAG AGCTCTAACC AGCTGAGCTA | 18960 |
| AAGGTCGGAC AAGATCATTA TAGCGGCGAA GGGGATCGAA CCCCCGACCT CCCGGGTATG | 19020 |
| AACCGGACGC TCTAGCCAGC TGAGCTACAC CGCCATGAAT CGGGAAGACA GGATTCTGAAC | 19080 |
| CTGCGACACC TTGGTCCCAA ACCAAGTACT CTACCAAGCT GAGCTACTTC CCGAGTTAAA | 19140 |
| TAGAAAAATG CACCCTAGAG GAGTCGAACC TCTAACCGCC TGATTCGTAG TCAGGTA CTC | 19200 |
| TATCCAGTTG AGCTAAGGGT GCTCCATATT ATGCCGAGGA CCGGAATCGA ACCGGTACGA | 19260 |
| TCGTTACCAA TCGCAGGATT TTAAGTCTCG TGCGTCTGCC AGTTCCGCCA CCCCCGCCCTC | 19320 |
| TCTAAGCGAA CGACGGGATT CGAACCCGCG ACCCCACCT TGGCAAGGTG GTGTTCTACC | 19380 |
| ACTGAACTAC GTTCGCACTG TTTTCTTCTA TCTAAAAATG CCGGCTACAT GACTTGAACA | 19440 |
| CGCGACCCCTC TGATTACAAA TCAGATGCTC TACCAACTGA GCTAAGCCGG CTCATTTGTT | 19500 |
| ATATCTTAAT GCGGGTTAAG GGACTTGAAC CCCCACGCCG TTAAGCGCCA GATCCTAAAT | 19560 |
| CTGGTGCGTC TGCCAAATTC GCCAAACCCG CATATATGAC CCGTACTGGG CTCGAACCAG | 19620 |
| TGACCCATTG ATTAAAAGTC AATTGCTCTA CCAACTGAGC TAACGAGTCT AAAATAACTT | 19680 |
| GCGTTACCTT AAACGGTCCG ACGGAATCGA CCCGGTAC | 19718 |

(2) INFORMATION FOR SEQ ID NO: 100:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 4117 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 100:

| | |
|--|-----|
| CCGTGGAAAA GTCTGGATAG TGAATGGTCT TCACACAATG ACCTGAAAGA AGCCTGAGAA | 60 |
| TAATTATGGA GAGTAGCATT CTGAGAGGTG TTAGCAGAAC CATATGACAG AGCTGTTTGA | 120 |
| AGAGGGAATA TTGAGGAGAA AAATCCTGAG CCTACCAGTT GGAGTTGGAA AGAGCTGACT | 180 |
| GT TAGATCAT GGTTTATTAT CCACAACCTG TGGATAACTT TGTGAATAAG AGAAGTTGCT | 240 |
| AAAGAAGGAG ATATATAACG ATGAAGAAAA TCAAACCGCA TGGACCGTTA CCAAGTCAGA | 300 |
| CTCAGCTAGC TTATCTGGGA GATGAACTAG CAGCTTTTAT CCACTTCGGT CCTAATACCT | 360 |
| TTTATGACCA AGAATGGGGG ACTGGACAGG AGGATCCTGA GCGCTTTAAC CCGAGTCAGT | 420 |

763

| | |
|--|------|
| TGGATGCGCG TCAGTGGGTT CGTGTGCTCA AGGAAACGGG CTTCAAAAAG TTGATTTTGG | 480 |
| TGGTCAAGCA CCACGATGGC TTTGTCCTTT ATCCGACAGC TCACACAGAT TATTCGGTTA | 540 |
| AGGTCAGTCC TTGGAGGAGA GGAAAGGGCG ACTTGCTCCT TGAAGTATCC CAAGCTGCCA | 600 |
| CAGAGTTTGA TATGGATATG GGGGTCTACC TGTCACCGTG GGATGCCCAT AGTCCCCTCT | 660 |
| ATCATGTGGA CCGAGAAGCG GACTACAATG CCTATTATCT GGCTCAGTTG AAGGAAATCT | 720 |
| TATCAAATCC TAACTATGGG AATGCTGGTA AGTTCGCTGA GGT'TTGGATG GATGGTGCCA | 780 |
| GAGGAGAGGG CGCGCAAAAG GTTAATTATG AATTTGAAAA ATGGTTTGAA ACCATTTCGTG | 840 |
| ACCTGCAGGG CGATTGCTTG ATTTTTCCTCA CAGAAGGCAC CAGTATCCGC TGGATTGGCA | 900 |
| ATGAACGAGG GSTATGCAGGT GATCCACTGT GGCAAAAGGT GAATCCTGAT AAACTAGGAA | 960 |
| CAGAAGCAGA GCTGAATAT CTTCAGCACG GGGATCCCTC GGGCACGATT TTTTCAATCG | 1020 |
| GAGAGGCAGA TGTTTCCATC CGTCCAGGCT GGTCTACCA TGAGGATCAG GATCCTAAGT | 1080 |
| CTCTCGAGGA GTTGGTCGAA ATCTACTTTC ACTCAGTAGG GCGAGGAACT CCACTCTTGC | 1140 |
| TTAATATTCC GCCGAATCAA GCTGGGCTCT TTGATGCAAA GGATATTGAA CGACTTTATG | 1200 |
| AATTTGCGAC CTATCGCAAT GAGCTCTATA AAGAAGATTT GGCTCTGGGA GCTGAGGTAT | 1260 |
| CTGGTCCAGC TCTTCCGCA GACTTTGCTT GTCGCCATTT GACAGACGGC CTTGAGACCA | 1320 |
| GCTCTTGGGC AAGCGATGCA GACTTGCCCA TCCAGTTAGA ACTCGACTTA GGTTCCTCTA | 1380 |
| AACTTTTGA TGTAATTGAG TTAAGAGAAG ATTTGAAGCT AGGGCAACGA ATCGCTGCTT | 1440 |
| TTCATGTGCA AGTAGAGGTG GATGGTGTCT GGCAGGAGTT TGGTTCGGGT CATACTGTTG | 1500 |
| GTACAAACG TCTCTTACGA GGAGCAGTTG TTGAGGCACA GAAGATACGT GTAGTCATTA | 1560 |
| CAGAATCACA GGCTTTGCCT TTGTTGACCA AGATTTCCTT TTATAAAACT CCTGGATTAT | 1620 |
| CAAAAAAGA AGTTGTTTCA GAACTAGCAT TTGCAGAAAA AAGCCTAGCT GTGGCAAAGG | 1680 |
| GAGAAAATGC CTATTTTACA GTTAAGCGCA GAGAATGTAG TGGTCCTTTA GAAGCTAAGA | 1740 |
| TTTCGATTCA ACCGGGGACA GGTGTCCATG GTGTGCGCTA TCAGGATGAG ATTCAAGTCC | 1800 |
| TTGCGTTTCA AACTGGTGAG ACTGAAAAAA GTCTGACGCT ACCAACCTTG TATTTCGCAG | 1860 |
| GAGATAAAAC CTTGGATTTC TATCTGAACC TAACGGTGGA TGCTCAGCTT GTGGATCAAC | 1920 |
| TTCAAGTCCA AGTTTCATAA AAGAAGAACC TTTGCGCGAT GCAAAGGTTT TTTTGGTTAT | 1980 |
| TAGTGACTTG GTAACCAGCT GAGGGTGAAA GTTAGTTGTT CAGCTTTTAA GAGGTCTTGG | 2040 |
| TGTTGAATAG TTGATACGAG TGTTTTGTCC AGTCGGCATT CTTTGACAAA GTTAAATGG | 2100 |
| TTGTGGTTTT GTTTAGTATG GATATCCAGC CATTTATCTT CTTTAGCGAG GTAGACTCGT | 2160 |

764

| | |
|--|------|
| AGATGGTCAA AGAGAGGGAT TCCGAGGTCA TAGCTTGGTT TTCCTGGACA GGTGGATAA | 2220 |
| AATCCGAGAG CTGACCAGAT GTACCAAGCA GAGAGACTAC CATTGTCTTC ATCTCCAGGA | 2280 |
| TAGGCTTCCC AACTTGGGTG AAAAGCTTTC TGACGGAGCG TCTTGATAAG AAGGGCAGTG | 2340 |
| TAGTCAGGGT AATCGCTGTA ACGGAAGAGA TAAGGAATGT GGAAACTAGG CTGGTTGGAA | 2400 |
| ATGGCTATTT GTCCAAAAGG AGCAGTAGCC ATCTCGCTCA TTTCGTGAAT TTCGTAACCA | 2460 |
| TAGCCTGTTG TTTCAAAGAG GGGAGCATCT TGACAGGCTT TCAAAAGATA GTTGCTAAAG | 2520 |
| GTTTCTTTTC CACCCATCAG TTGGATTAAAG CCAGGGATGT CGTGAGAAC GCCTAAAGTA | 2580 |
| GCTTGAATGG CAGAGCATTC AGCGTAGTCT CGCCCCAAC TATAAGGAGA GAAGTCAGGG | 2640 |
| TGAAAGTTTC CTTGATTGTC TCGTGCTCGC ATGTAACCTG TCTCAGCGTC AAATAGCTGG | 2700 |
| CGGTAATTTT GTGAAGCAGC CTTGTAGGTT TCAGCGATTT CTATGTTCTC TAGTTTTTTG | 2760 |
| GCACAGCTGG CGATACAAA GTCACATATAG GCATAGTCTA GAGTATGGCT AACACTTTCG | 2820 |
| TGGTGGTCGG TAGAGAGGTA ACCTAGTTCT TGGTATTGGG CTAGTCCGTG GCGGCCATTG | 2880 |
| ATGCCGAGAG GGTCCGCTTT GCTGGCTGTT TCGAGCATGG CTTGGAAGAG TTCTCCTCT | 2940 |
| AGGTCGGGG TCATGTCCTT GCAGGCGCTA TCTGCGATAA TACCGTCTAA AAGTGACCT | 3000 |
| GGCATCATAC CCCGTTTCATC TGGAGCCAGC CATTTTGGAA GGAAACCAGT ATCGCGGTAG | 3060 |
| CTATTGAGGA AACCTTCTAA AAAGCGTTGA TAGTGCTCCG GTATGATAAG GGCAAAGAGG | 3120 |
| GGGAAGGTGG TGCGGAAGGT ATCCAGAAA CCATTGTTGC TAAAGAGGAC ACCAGGCTTG | 3180 |
| ACAGTACCAG TAGCCAGATC CATGTGGATG GCTTGCCCTG ATTCATTAAT CTCATAAAAA | 3240 |
| GTCTGTGGGA AGAGGAAGAG TCTGTAGAGG CAGTGGTCAA AGAAGGTTG GTCAGCCTCT | 3300 |
| CCTGTCTCTA TAATGTCAAA ACGATGGAGG AGATTTTCCC AATCCACTTG GGCACCTGAT | 3360 |
| TTACAGCTAT CAAAATCTTC TTGAGGTAGA TTGATTAGAG CTTGAGAAGG AGAGATGAAA | 3420 |
| GAAGTGGCTA GTTGCACTC GGTGTGACTA CTTGCTAAGT CAATTCGCCA GTCTCCAGCT | 3480 |
| TCTTGGCTGA TAGCAAGAAT ATCCGTGTTT ATTTGCAGGG CAGTGAACAT CGTTAGCGAA | 3540 |
| TTTTTGTTAG TTTCACTTTT ACCTTCTTGT CGCAGGGCAA GAGTCCGCTT ATCTACTTGC | 3600 |
| TCTACTGTCA GTTCATCTGC TGCCTGAAGA TAGAGGGAGA GGGCTTTGCC TTGCTTTTGA | 3660 |
| TTCAAACGAA TAGAAGCACC ATAGCAAGTC GGTGTGAGCT GGGTTTCAAT CTGATAACGC | 3720 |
| AGAGAAAAGA GCTTCAAATA GTGAGGCTGG AAGCAAGCTT TATCTATATC ATAAGAAGAC | 3780 |
| TGGCGGTGAA AGAGGCTGTC TCCCCCAGT TGACTGGTGA CAGGTGTCAG AAGGAGCCAA | 3840 |
| GAGTAGTCCC CAATCCAAGG ACTGGGCTGG TGAGTTAATC GAATCCCCTG AAAGATAGGC | 3900 |
| AGATGTGGAT CAAAAACCA AGATCCATCC TGGTCACTGG TCTGGGGCAC AAAGTAATTC | 3960 |

765

| | |
|---|------|
| ATCCCAAAAG GCACGCCTGT GTATGGCAGG GTATTTCCCC GAGAAAAGGC ATGCTTGTTC | 4020 |
| GTAGTTCCAA AACGGGTATC GATGGTATCA AGTAGTGGTT TCATAGTCTT TCCTTTAGCT | 4080 |
| GTTTTCTAC ATTATATCAG TAATAGAGGG CCTTTAG | 4117 |

(2) INFORMATION FOR SEQ ID NO: 101:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2727 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 101:

| | |
|---|------|
| CTGGTTCAAT TATTATTCAC TCTAAGTAGT CATATGTTCT TTATTTATGT GAGTTTTTAC | 60 |
| CTTTTAAAGG ATCTTGTTAG ATGGGAGAAG GTTTTAAAAG TGACAGATGA TAATACAAGA | 120 |
| AAAGTTCGTT TATTAGTAGC CTTTTTTAGC ATTGTCATAG GCTACATCCT GAGTTCTTTC | 180 |
| TTTATTAGCC TGTATCATTT GTGGCAAGAA GCGCTTAGAG GATTATTATG AAATCAAGAG | 240 |
| TAAAGGAAAC GAGTATGGAT AAAATTGTGG TTCAAGGTGG CGATAATCGT CTGGTAGGAA | 300 |
| GCGTGACGAT CGAGGGAGCA AAAAATGCAG TCTTACCCTT GTTGGCAGCG ACTATTCTAG | 360 |
| CAAGTGAAGG AAAGACCGTC TTGCAGAATG TTCCGATTTT GTCGGATGTC TTTATTATGA | 420 |
| ATCAGGTAGT TGGTGGTTTG AATGCCAAGG TTGACTTTGA TGAGGAAGCT CATCTTGTC | 480 |
| AGGTGGATGC TACTGGCGAC ATCACTGAGG AAGCCCTTA CAAGTATGTC AGCAAGATGC | 540 |
| GCGCCTCCAT CGTTGTATTA GGGCCAATCC TTGCCCCTGT GGGTCATGCC AAGGTATCCA | 600 |
| TGCCAGGTGG TTGTACGATT GGTAGCCGTC CTATTGATCT TCATTTGAAA GGTCTGGAAG | 660 |
| CTATGGGGGT TAAGATTAGT CAGACAGCTG GTTACATCGA AGCCAAGGCA GAACGCTTGC | 720 |
| ATGGTGCTCA TATCTATATG GACTTTCCAA GTGTTGGTGC AACGCAGAAC TTGATGATGG | 780 |
| CAGCGACTCT GGCTGATGGG GTGACAGTGA TTGAGAATGC TGCGCGTGAG CCTGAGATTG | 840 |
| TTGACTTAGC CATTCTCCTT AATGAAATGG GAGCCAAGGT CAAAGGTGCT GGTACAGAGA | 900 |
| CTATAACCAT TACTGGTGTT GAGAACTTC ATGGTACGAC TCACAATGTA GTCCAAGACC | 960 |
| GTATCGAAGC AGGAACCTTT ATGGTAGCTG CTGCCATGAC TGGTGGTGAT GTCTTGATTC | 1020 |
| GAGACGCTGT CTGGGAGCAC AACCGTCCCT TGATTGCCAA GTTACTTGAA ATGGGTGTTG | 1080 |
| AAGTAATGA AGAAGACGAA GGAATTCGTG TTCGTTCTCA ACTAGAAAAT CTAAAAGCTG | 1140 |
| TTCATGTGAA AACCTTGCCC CACCCAGGAT TTCCAACAGA TATGCAGGCT CAATTTACAG | 1200 |

766

| | |
|--|------|
| CCTTGATGAC AGTTGCAAAA GGCGAATCAA CCATGGTGGG GACAGTTTTC GAAAATCGTT | 1260 |
| TCCAACACCT AGAAGAGATG CGCCGCATGG GCTTGCATTC TGAGATTATC CGTGATACAG | 1320 |
| CTCGTATTGT TGGTGGACAG CCTTTCAGG GAGCAGAAGT TCTTTCAACT GACCTTCGTG | 1380 |
| CCAGTGCGGC CTTGATTTTG ACAGGTTTGG TAGCACAGGG AGAAACTGTG GTCGGTAAAT | 1440 |
| TGGTTCACCT GGATAGAGGT TACTACGGT TCCATGAGAA GTTGGCGCAG CTAGGTGCTA | 1500 |
| AGATTACAGG GATTGAGGCA AGTGATGAAG ATGAATAAGA AATCAAGCTA CGTAGTCAAG | 1560 |
| CGTTTACTTT TAGTCATCAT AGTACTGATT TTAGGTACTC TGGCTCTAGG AATCGGTTTA | 1620 |
| ATGGTAGGTT ATGGAATCTT GGGCAAGGGT CAAGATCCAT GGGCTATCCT GTCTCCAGCA | 1680 |
| AAATGGCAGG AATTGATTCA TAAATTTACA GGAAATTAGG CTGGAGAACC AGCCTTTTTC | 1740 |
| TAAAGATAAG GAGAAATATG AACAAAAAAA CAAGACAGAC ACTAATCGGA CTGCTAGTGT | 1800 |
| TATTGCTTTT GTCTACAGGG AGCTATTATA TCAAGCAGAT GCCGTCGGCA CCTAATAGTC | 1860 |
| CCAAAACCAA TCTTAGTCAG AAAAAACAAG CGTCTGAAGC TCCTAGTCAA GCATTGGCAG | 1920 |
| AGAGTGCTTT AACAGACGCA GTCAAGAGTC AAATAAAGGG GAGTCTGGAG TGAATGGCT | 1980 |
| CAGGTGCTTT TATCGTCAAT GGAATAAAAA CAAATCTAGA TGCCAAGGTT TCAAGTAAGC | 2040 |
| CCTACGCTGA CAATAAAACA AAGACAGTGG GCAAGGAAAC TGTTCCAACC GTAGCTAATG | 2100 |
| CCCTCTTGTC TAAGGCCACT CGTCAGTACA AGAATCGTAA AGAACTGGG AATGGTTCAA | 2160 |
| CTTCTTGGAC TCCTCCAGGT TGGCATCAGG TCAAGAATCT AAAGGGCTCT TATACCCATG | 2220 |
| CAGTCGATAG AGGTCATTG TTAGGCTATG CCTTAATCGG TGGTTTGGAT GGTTTTGATG | 2280 |
| CCTCAACAAG CAATCCTAAA AACATTGCTG TTCAGACAGC CTGGGCAAAAT CAGGCACAAG | 2340 |
| CCGAGTATTC GACTGGTCAA AACTACTATG AAAGCAAGGT GCGTAAAGCC TTGGACCAAA | 2400 |
| ACAAGCGTGT CCGTTACCGT GTAACCCCTT ACTACGCTTC AAACGAGGAT TTAGTCCCT | 2460 |
| CAGCTTCACA GATTGAAGCC AAGTCTTCGG ATGGAGAATT GGAATTCAAT GTTCTAGTTC | 2520 |
| CCAATGTTCA AAAGGGACTT CAACTGGATT ACCGAACTGG AGAAGTAACT GTAACCTCAGT | 2580 |
| AAAAGATACG CCTACACTCC TATGTCACTT ATGGATGTAG GAGTTCCTTT TACTAGTTTA | 2640 |
| AGCAGGACTA AGACAGGTAC TAAGACAAAA TAGCAACTTC TAAAACTAAC TTCCAGTTTT | 2700 |
| GGGAGAGAGA TGAAGTTAC TTTGAGA | 2727 |

(2) INFORMATION FOR SEQ ID NO: 102:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 5717 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

767

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 102:

| | |
|--|------|
| TTTTTTGTAG ATTTAAGTGG GGTGCAATTC CTAAAAATA AAAACAATT TTTGAAAATT | 60 |
| ATGTTAGCAG GAATTGCTTC AAATTCGATT TTATCACTTA CAGGTTTACT TGTTTTATTG | 120 |
| TTCACATCGT ATAAATTGCT TGGACTCTTA TTTTATCA TTAACCTAGG TATGATTTT | 180 |
| ATTAATTCAA TTCCTTTTTT TCAGTATGAT AGTGGTATTA TTTTAAGATA CTTGAATTCT | 240 |
| AACAATAATA ACTTGAATTT TCAATATATA GTTCAACTTT TAATAGCATT TGTATTATT | 300 |
| TATTTTCCTT TGAGTCAACT ATTACAGTTT TTGACACCA ATATTATTGT TCGTAGTATA | 360 |
| GGAGGGGTGG TTGTTTCTAT ACTGCTTCT ATATTATATA TGATAGGAAG GACGAAATAT | 420 |
| GTTCTACGTA AATAGTTATG TTTTGTCTTA TAAAAAGAA GGTATAATGT ATTTACGTGG | 480 |
| TCGGAGTATG CGGGAATAG CTATAGAACC TCAAAATTCG CAAGAATTTA TCAACGATCT | 540 |
| ATTTAATAGT TGTAAGGAAC TATTAGAGAT AGAAGAAGTA TTAGGCAGTA AACTAACATT | 600 |
| TGAACATAAA ATGAACAAAT TTTAATTTCG GATGAGATAG ATATTGATAG TAGATATTCT | 660 |
| AGAACTAAAG GTTACTATTC GTTATTTTAT AATGAAGACT ATAATAAAAT ACAGAATAAA | 720 |
| ACAGTATTAG TATTAGGAGC AGGAGTCTTA GGATGTTATA TATCTCTAAG TCTAAGTATG | 780 |
| TATGGAGTGA GGAACTTAT TGTCGCTGAT TACGATATAA TAGAACCATC AAATTTAAAT | 840 |
| AGGCAATTC TTTATACAGA GTCGGATGTT GGTAAGGAGA AGATTAATGT TCTTCTGAA | 900 |
| AAAATACACA AGTATAATTC AGATGTCAG GTAGTACCTA TTTCTATTAA AGTTTCTTCA | 960 |
| GTAGAAGAAT TAGAAAAAT TGTTGCGGAA TATGGGAGTA TAGATTTTAT CGTTAAAGCA | 1020 |
| ATTGATACGC CCATTGATAT TATAAAAATT GTCAATCAAT TTGCTGTATC GCATAAGATA | 1080 |
| TCCTACATAT CAGGAGGGTT TAATGGATGC TATCTTATTA TTGATAATAT ATATATCCCT | 1140 |
| ACCATCGGTT CTGCTTTGG TTGTCGGAAT ATAAACAAAG ATATAAATAA GTACACTTTA | 1200 |
| TCTGATAAGA CAAAGTGGCC GACTACACCA GAGATGCCTG CTATTTTGGG AGGGATAATG | 1260 |
| ACTAATTTAA TAATTAAAT ATTTCTGGGA TGTATAATG AAATCCTAAT AGATAACGCT | 1320 |
| TACGTTTATA ATATGAGAAA TCATGCTCTA AGTCAAGAAA AATATGTTCT GGAAAACGGA | 1380 |
| GAATGTCCAA TTTGTAAAAA AATAATAAAG TGAAAGATAA CAATATTAGA GCGAAAACAT | 1440 |
| TTATTCGTTT AGTTTGTTTT TGCTTATTAT CAGGAGGAGT AGCTTTTTTA TCTGCTATTG | 1500 |
| GGCAGTTCAC TGTATAGAA ACACAATTAA TAGTATTGTT CTTGGGTATT ATTTTGTCTA | 1560 |
| TATATTATGC TTAATACAAAT AAAAATATTC AAACATCATT GGAAAATATA GTATGGCTTT | 1620 |

768

| | |
|---|------|
| TTTCATCGTT TGAGATTPTA TTTTGTCTTG TTAATTTTAG AACATTATT CAGTTACCAG | 1680 |
| TGGATATTTT TATTGGTATG ATAATATTTT TAATGCTGTG GATATTTATT ATGTTAGGTA | 1740 |
| TAGTGTGTCT TAGTTATTAT ATAACTTTAT TATTTAGCAA GGAGGCTTAG TATGTTTAAA | 1800 |
| AAAATAGGTA TAATGAGCAT TTGCATATAT ATAATTATTT TATACTGCTT GAGAATGTAT | 1860 |
| CGTATTATCA ATAATATTGA AACAACTCTG CTAACGGTTA TATGCTTAAT GTTATTGTTT | 1920 |
| TTTTTAAGAC GTTTATTTGA TAAAGATAAG TAAATAGATG TTAAGTAAAA ATGTAGAATA | 1980 |
| TAAAGGAGGT GCAATGAGTA TGATTGAAGT TAGCCATTTA TCAAAAAGTT TTGGTGATAA | 2040 |
| AATAGCTTTA AATAATATAA GCTTCACTGT TAAAGAAGGT TAGATTTTTG GATTTTTAGA | 2100 |
| ACCATCTGGT TCTGGAAGA CCACAACGAT TAATATTCTG ACTGGGCAGT TCCTTGCCGA | 2160 |
| TAAAGGACAA TCTATTATTT TGGGACAAAA ATCTCAAAAT TTAACAAGCG GTGAATTAAA | 2220 |
| GAGAATTGGA TTGGTTAGCG ATACAAGTGG ATTTTATGAG AAAATGTCTC TGTATAACAA | 2280 |
| TCTTCTTTT TATAGTAAAT TTTATAATAT TAGTAAATCA CGTGTTGATA ATTTGTTAAA | 2340 |
| GCGAGTAGGA TTATATGATA GTCGCAAGAT GGTAGCAGGA AAATTATCCA CTGGAATGAG | 2400 |
| GCAACGAATG CTTTTAGCAC GAGCTCTTAT CAACAACCCC GCTGTACTCT TTCTGGATGA | 2460 |
| ACCGACCTCA GGTCTAGATC CCACAACCTC TCGAACAATT CATGAGTTAA TTTTAGAATT | 2520 |
| GAAAACAGCA GGGACAACGA TTTTCTAAC GACTCATGAT ATGAATGAAG CAACTCTTTT | 2580 |
| ATGTGATTAT GTTGCCTTAT TAAATAAAGG GAAATTAGTT GAGCAAGGAG CTCCTTCTGA | 2640 |
| ACTCATTCAA AGATATAATA AAGATAAAAA GATTAAGGTT ACAGATTATA ATGGGAATCA | 2700 |
| GATAACTTTT GATTTTACAT CACTAGAACA GGTATCTCAG ACTGATCTGG AAAATATTTT | 2760 |
| TTCAATTCAT TCATGTGAGC CTACTTTAGA AGATATTTTT ATCACATTAA CAGGAGGAAA | 2820 |
| GCTAAATGCT TAAACGGTTT CTGGCTTTGG TATGGTTGCG TTGTCAAATC ATCCTTTCCA | 2880 |
| ATAAGAGTAT TTTATTGCAA GTTTTAGTGC CTTTGTCTTT CACATATTTT TATAAATATC | 2940 |
| TTATGGAAAC ACAGGGAAG GTCAACGATC AACAGGCATT AGTTCCTTTG ATGATGTGTT | 3000 |
| TACCTTTTTC TTTTCTTTG GCTGTTGGAA GTCCTATAAC TATTATCTTG TCTGAAGAAA | 3060 |
| AAGAAAAGTA CAATTTACAA ACTCTTCTGT TGAGTGGTGT TAAAGGCTCC GAATACATTT | 3120 |
| TATCAACTAT GTTCTTCCT TTTTGTCTAA CTTTGTGAT TATGGGAACT ACTCCTCTTA | 3180 |
| TTTTAGGAGT TACAATTGTA CATACTTTTA ATTATATTAC AATCGTCTTT CTAACCTCTT | 3240 |
| TATCCATCAT TTTATCTTAT TTATTGATAG GTTTAACCGC GAAGAGCCAA GTAGTAGCTC | 3300 |
| AGGTATCAG TCTTCTGCT ATGATTTTAG TTGCTTCTTT ACCGATGCTA TCTGGTTTGG | 3360 |
| ATAAGACAGT TGCGAAGATA ACAGATTATA GTTTTATGGG ACTATTTACT AAGTTTTTCA | 3420 |

769

| | |
|--|------|
| CAAAATGGGA GGAATTTTCA TGGAATAAAA CTCTAATTC TAATCTAACA CTACTTATTT | 3480 |
| GGATTGTTCT TCTATTAAC TTAATTACGA TAACTATTAG GAAAAAGAAA ATTTCTTAAT | 3540 |
| TGAGTTATTT TAATGATTAT AAACACAAGT GGAAGGAAA AAATGAAGT ATCTTTTGA | 3600 |
| CAGCAATTCT ACAGAATAGT CTTATTGCTA TATTTTGATT TGAGTGACG AAAAAAGAAA | 3660 |
| AATAACAATA GTGCTCATA TAATTGCAGA AGTTTGGGT GATAAGATAA CTGATAAATT | 3720 |
| GCAATAAAAA ATGCAACATT TTAAATCTC CTCTATAAGT GCTTCAAAA GTGCTTCAAA | 3780 |
| ACCTGTCTTG TAATCCAAGT ATTTTGGGG ACGGTGATTA ATAAGCTAGC AAAGCATCAT | 3840 |
| TAAGGATTTT TTCGGTAATT GTTGCCAAAT CGGTTTAA GAATACTCAC GAAGAAGTCC | 3900 |
| ATTCGCATTC TCATTACTTC CCCTTTGCCA AGATGAATAG GCATCCGCA AATAAACAG | 3960 |
| AATTCACATT TGTCAATTA AAGGTAACA AGCAACTCT TTTCTCTGT CCGAAGTGAA | 4020 |
| AGTCTTAACT TATCTTTTG GAAAGAGTCT TGTGAGGTGT TCAATAGCAG TCAACATGGA | 4080 |
| TTTAGCTGTT TTTACTTGAC AAGTGCTAGT AGAAATAATA GAATAGTAAA AAACCTTTAA | 4140 |
| AGCAGTCCAG AGAGGCAGCT AAGGTTAGAC GGTGAAAGG TGGAGACTAC CCATTTTTCG | 4200 |
| TGGAACCTTG CTGTTGGCAG GTTCCTTTT TCGTGGCTTC TGTGGCCAG ACTCTCTAC | 4260 |
| TAGTAAAGGT AAAAGGAGAA ACCTATGCGA GAACATCGTC CAATCATTC TCTTGATTTT | 4320 |
| CCTAGTTTTC AGGCGGTCAA GGAATTTTTC GCTCTTTTC CAGCAGAAGA AAGCCTTTAT | 4380 |
| CTCAAGGTAG GGATGGAGCT TTATTACGCA GCGGGGCTG AGATTGTGTC CTACTTAAAA | 4440 |
| GGTTTGGGTC ATAGTGTCTT TTTGGATCTC AAACCTCATG ACATTCTTAA TACAGTCAAG | 4500 |
| TCAGCCATGA AGATCTTGTC TCAGCTTGGT GTCGATATGA CTAATGTCCA TCGGCTGGT | 4560 |
| GGTGTAGAGA TGATGAAGGC GCGCGTGAA GGTCTTGGGA GTCAAGCCAA ATTGATCGCT | 4620 |
| GTAACCTCAGC TCACATCAAC GTCAGAAGCT CAGATGCAGG AGTTTCAAAA TATCCAAACC | 4680 |
| AGTCTGCAAG AGTCTGTGAT TCACTATGCC AAGAAGACAG CTGAAGCTGG CTTGGATGGT | 4740 |
| GTGTTTGGCT CGGCTCAGGA AGTACAAGTC ATCAAGCAGG CTACCAATCC AGATTTTATC | 4800 |
| TGTCTGACAC CAGGGATTTC TCCAGCTGGT GTTGCAAGT GAGATCAAAA ACGAGTCATG | 4860 |
| ACACCTGCTG ATGCTATCA AATCGGCAGT GACTATATCG TAGTGGGACG TCCCATACC | 4920 |
| CAAGCTGAGG ATCCTGTGTC AGCTTATCAT GCCATCAAG ATGAATGGAC ACAGGACTGG | 4980 |
| AATTAAAGAA CTAGATTAGA AAAATAAAG GAGAATACCA TGACACTTGC TAAAGATATC | 5040 |
| GCTAGCCACC TCTTGAAAT CCAAGCCGTT TACCTCAAAC CAGAGGAACC CTTCACTTGG | 5100 |
| GCATCTGGTA TCAAGTCACC GATTTACACT GATAATCGTG TGACACTAGC CTATCCAGAA | 5160 |

770

| | |
|---|------|
| ACTCGTACCC TAATTGAAAA TGGTTTGTG GAAGCTATCA AAGAAGCCTT TCCTGAAGTA | 5220 |
| GAAGTGATTG CAGGAACTGC AACAGCAGGG ATTCCACACG GAGCCATTAT TGCTGATAAG | 5280 |
| ATGGACTTGC CTTTTGCCTA CATCCGTAGT AAACCAAAAG ACCACGGAGC TGGTAATCAA | 5340 |
| ATCGAAGGTC GCGTAGCTCA AGGTCAAAAA ATGGTAGTGG TTGAAGACCT TATTTCAACG | 5400 |
| GGTGGTTTCTG TTCTTGAAGC TGTAGCAGCA GCCAAGCGAG AAGGAGCAGA TGTACTTGA | 5460 |
| GTTGTAGCGA TTTTCAGCTA CCAATTGCCA AAAGCAGATA AGAAGTTTGC AGATGCTGGT | 5520 |
| GTTAAACTTG TGACGCTTTC AAATATAGC GAGCTTATCC ATCTAGCCCA AGAAGAAGGT | 5580 |
| TACATCACGC CAGAGGGCCT TGATCTTCTA AAACGCTTTA AAGAAGACCA AGAAAATTGG | 5640 |
| CAAGAAGGTT AGGTCAGTAA GATAAAGAGA GACGAGGCTA CCGAGTCTCT TTTACCATTT | 5700 |
| TATTTAAAAT ATGACAG | 5717 |

(2) INFORMATION FOR SEQ ID NO: 103:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 5558 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 103:

| | |
|---|-----|
| CCTGGACTTT CTAAATGAA ATCTTGCAC CTGGATCAAG CCCTTCATGA GCATTTTTCA | 60 |
| GAAGAAGAAT TAGCTGGTCA CTTTCATGTC CTTCTATGGA CTTTTTTTAC AATGGCATTG | 120 |
| CTATCACACC CAATACCTAT CTAAGCGCCT GGTCGTAAA CTTTATGCA GCTCTTCCTC | 180 |
| TAAATTTCTT AATTGTTGAA CCAATTGCCC GTTTTATACT AAGTCTTTT CAGAAACCAT | 240 |
| TTACTGGGGA AGAAGTTGAA GATTTTCAAG ATGATGATGA AATCCCACT ATTATCTAAG | 300 |
| CCAGTTCTGT AAACACTAA TATTTGAAAT CCACTTCCTT TTAGGGTGCA ATGGTTATAA | 360 |
| ATGAATTTTT GAGAGGATCA GAATGAAAA ACTAGCAACC CTTCTTTTAC TGTCTACTGT | 420 |
| AGCCCTAGCT GGGTGTAGCA GCGTCCAACG CAGTCTGCGT GGTGATGATT ATGTTGATTC | 480 |
| CAGTCTTGCT GCTGAAGAAA GTTCCAAAGT AGCTGCCCAA TCTGCCAAGG AGTTAAACGA | 540 |
| TGCTTTAACA AACGAAAACG CCAATTTCCC ACAACTATCT AAGGAAGTTG CTGAAGATGA | 600 |
| AGCCGAAGTG ATTTTCCACA CAAGCCAAGG TGATATTCGC ATTAACTCT TCCCTAACT | 660 |
| CGCTCCTCTA GCGGTTGAAA ATTTCTCTAC TCACGCCAAA GAAGGCTACT ATAACGGTAT | 720 |
| TACCTTCCAC CGTGTCATCG ATGGCTTTAT GTTCCAACT GGAGATCCAA AAGGGGACGG | 780 |
| TACAGGTGGT CAGTCCATCT GGCATGACAA GGATAAGACT AAAGACAAAG GAACTGGTTT | 840 |

771

| | |
|---|------|
| CAAGAACGAG ATTACTCCTT ATTTGTATAA CATCCGTGGT GCTCTTGCTA TGGCTAATAC | 900 |
| TGGTCAACCA AACACCAATG GCAGCCAGTT CTTTCATCAAC CAAAACCTCTA CAGATACCTC | 960 |
| TTCTAAACTC CCTACAAGCA AGTATCCACA GAAAATTATT GAAGCCTACA AAGAAGGTGG | 1020 |
| AAACCCTAGT CTAGATGGCA AACACCCAGT CTTTGGTCAA GTGATTGACG GTATGGATGT | 1080 |
| TGTGGATAAG ATTGCTAAGG CCGAAAAAGA TGAAAAAGAC AAGCCAACTA CTGCTATCAC | 1140 |
| AATCGACAGC ATCGAAGTGG TGAAAGACTA CGATTTTAAA TCTTAAAAAC CAAAAAATA | 1200 |
| CAGTATCCAC ATTTCGGTACT GTATTTCTTT TACTCTCATT CTTAAGTTAA ATTATTAAAA | 1260 |
| TCCCATATTT GGTCTATCCA GCCTTCATAA AAGTCTGGCT CGTGGCAGAC CATAAGGATA | 1320 |
| GATCCCCTAT ATTCTTTGAG AGCGCGTTTG AGCTCATCCT TTGCATCCAC ATCCAAATGG | 1380 |
| TTGGTCGGCT CGTCCAGCAC TAAAACGTTG TTTTCACGAT TCATCAAGAG ACAGAAACGA | 1440 |
| ACCTTGCGTT GCTCTCCCC TGATAATACT TGAATCTGGC TTTCAATATG TTTGGTTGTC | 1500 |
| AAACCACAAC GGGCAAGGCG TGCACGGACT TCTGCTTGAT TAAGGGCAGG AAAGGCATTC | 1560 |
| CAGACAGCTT CAAGAGGAGT TTGGCGATTA CCGCCTTCTA CTTCTGTGTC AAAATAACCA | 1620 |
| AGTTCTAAAT AATCTCCAGC CTCCACTTCC CCAGCGATTG GCGAGATAAT GCCCAAGAGA | 1680 |
| CTCTTCAAGA GAGTTGTTTT TCCAATACCA TTAGCACCAA TAATCGCAAC CTTTGTGATTG | 1740 |
| CGTTTGAAGG TAAGATTTAA AGGCTTAGTA AGAGGACGGT CGTAACCAAT TTGCAAGTTC | 1800 |
| TTGGCTTGGG AGATAAAGCG CCCTGGTGTA CGAGCTGGTT TGAAATCAAA GGATGGTTTT | 1860 |
| GGTTTCTCAC TTTGGAGTTC GATAATATCC ATCTTATCCA ATTTCTTTTG ACGAGACATA | 1920 |
| GCCATATTAC GAGTTGCAAC ACGGGCTTTA TTACGAGCCA CAAAGTCCTT GAGGTCTGCA | 1980 |
| ATCTCTTTCT GCTGGCGTTC GTAGGCTGCC TCTAGCTGAG ATTTCTTCAT AGCATAAACT | 2040 |
| TCTTGGAAct GGTAGTAGTC ACCAGAGTAA CGCGTCAGCT GTTGATTTTC CACATGATAG | 2100 |
| ACAATATTAA TAACGTCATT GAGGAATGGA ATATCGTGCG AAATGAGAAC AAAGGCATTC | 2160 |
| TCATAGTTTT GGAGATAGCG CTTGAGCCAA TCAATATGCT CAGCATCCAA GTAGTTGGTC | 2220 |
| GGCTCGTCCA ACAGCAAGAT ATCAGGCTTT TCAAGGAGAA GTTTTGCCAA AAGCACCTTG | 2280 |
| GTTCTTTGCC CACCTGACAA AGAAGTTACA TCCGTATCCA TGCCAAAGTC CATAACACCA | 2340 |
| AGAGCACGCG CTACTTCGTC AATCTTAGCA TCCAAGGTAT AGAAATCACG ACTCTCCAGA | 2400 |
| CGGTCTTGAA GTTCTCCTAC TTCTTCCATG AGAGCATCAA CATCCGCGCC GTCTTCAGCC | 2460 |
| ATTTTCATAT AGAGGTCATT GATACGAGCT TCAGCTTTGA AAAGCTCATC AAAAGCCGTA | 2520 |
| CGGAGAACAT CACGCACCGA CTGTCTTTCA GCAAGGACAG AGTGCTGATC CAAGTAACCA | 2580 |

772

| | | | | | | |
|-------------|------------|------------|------------|-------------|------------|------|
| GCCGTCACAT | ATTTGGACCA | CTCAACCTTT | CCTTCATCTG | GCAGCATTTT | ACCAGTCACG | 2640 |
| ATACTCATAA | AGGTTGATTT | TCCTTCACCA | TTGGCACCGA | CCAGGCCGAT | ATGTTCTCCC | 2700 |
| TTGAGGAGAC | GGAAGGACAC | ATCTTCAAAA | ATTGCACGGT | CACCAAAACC | GTGACTCAGA | 2760 |
| TTTTTAACTT | CTAAAATACT | CATTTTAATT | CCTTACCTTG | TTTTTATGTA | ATCGTTTATA | 2820 |
| AAGGAGCCAA | GCCAGATAGC | CACCCAAAGT | GTTGGTCCAC | AAATCATCAA | TCTCAAAGAC | 2880 |
| GCGATTGAAA | TCAAAGAAAA | AGTCCAAGAT | TAATTGCGTA | CACTCGATTG | CAAGACTCAC | 2940 |
| AAGAAAACATA | AAAAGAAGGA | CCTTTTTTGT | TTTCCGCAAA | TTTGGAAATA | GATAAAGGAG | 3000 |
| TTGGAAAATC | AGAGGAAAAA | ACAAGAAGAC | ATTGAGGATA | TTTTGTAAAA | AAATCCAACA | 3060 |
| TAATTGTCCA | ATGTCACTCA | CTTCGCCCAG | TTTCCAGAGA | GAATTGAAAG | GAGTCAAAAG | 3120 |
| AAAAACCAGG | CGTCCAAGAT | GCTGAATACC | TGGAGTTCCC | ACTCCCACGG | TAGATTGTTT | 3180 |
| TTGAGGAGTA | AAGCAAAAAC | AGACAATGCA | AATGCTATAG | AAAATGACTC | CCCAGACCRA | 3240 |
| AATATGATTA | TAAGTCTTCT | TCATCATTA | GGATTTACCG | CTGCGACTGC | CTTCTGGCGG | 3300 |
| TCACGTTTCA | TTGTGTTAGA | GCGCAATTGT | CCACAAGCTG | CGTCAATATC | TGTACCATGC | 3360 |
| TCTTGACGAA | CCACACAGTT | GACCCCTTTT | TTCTTAAGCG | TATCATAGAA | AGCCAACACG | 3420 |
| CACTCTTTGG | GACTACGGCT | ATATTGGTCA | TGCTCACTAA | CTGGGTTATA | AGGAATCAAG | 3480 |
| TTTACATAAG | ACAATTCTCT | GATGTTCTTG | AGCAATTGAG | TCAATTCCAA | GGCTTGTTCT | 3540 |
| ACACCGTCGT | TGACTTCATT | AAGCATGATA | TATTCAAAGG | TTACACGACG | GTTTGTGTC | 3600 |
| TCAATGTAGT | ATTCAATAGC | AGCAAAGAGT | TTTTCAATCG | GAAAGGCACG | GTTAATCTTC | 3660 |
| ATGATACTTG | AACGAAGTTC | ATTGTTAGGT | GCGTGAAGAG | ACACGGCAAG | ATTGACCTGA | 3720 |
| ACCCCTTCAT | CAGCAAAGTC | ACGAATTTTA | TGAGCCAAAC | CTGAGGTGTA | AACCGTGATG | 3780 |
| TGACGAGCAC | CGATAGCCAT | TCCTTTATCA | TCATTGATAG | TACGAAAGAA | ATTCAAGACA | 3840 |
| TTGTTGTAAT | TATCAAAGGG | CTCACCATT | CCCATGACAA | CGATATGGCT | GATGCGTTCA | 3900 |
| TCCTGACCAC | GCTCATCAAA | GTATTTCTGA | ACCAGCATGA | TTTGCGCTAC | GATTTACCCG | 3960 |
| TTATTGAGGT | CACGTTGCTT | CTTAATCAAA | CCAGAGGCAC | AGAAGGTACA | ACCGATATTA | 4020 |
| CAGCCGACCT | GAGTGGTCAC | ACAGACAGAT | AAACCATAGT | GTTGACGCAT | GAGTACAGTC | 4080 |
| TCAATTAACA | TACCGTCGGG | CAATTCAAAG | AGATATTTGA | CTGTACCATC | AGCAGACTCT | 4140 |
| TGCACAATAC | GTTGTTTCAA | GGGATTGACC | ACAAACTGGT | CATTGAGCTT | AGCAATCAAA | 4200 |
| TCCTTGAAA | GGTTGGTCAT | TTCTTCAAAT | GACTGCACAC | GTTTACGGTA | GAGCCATTCC | 4260 |
| CAGATTGAT | CTGCACGGAA | TTCTTTTCT | CCCTGCTCCA | ATACCCATTG | CTGCATGGTT | 4320 |
| TGATGTACCA | AACTATGAAT | TGAGGGTTTC | ATTTCTTCTC | CTTATTTCTCT | ACTCACTTCT | 4380 |

773

GACGAATGAC AAAATGACGT TGTCCCTTGT CGTCTTTCTG ACGACGTCTA TTTTCTTAT 4440
 CTGCATTCTGA CTTTCGTTTA GTTTGAGTCG GTTCTTTCC TTTCTAGAA GGTGTTTCTT 4500
 CTTCCTCTT ACGCATTTTC TTGTCAAATG ATGCTCGCTT AGGGGCTTCA TTTCTAAGA 4560
 CAAAATAGGC ACAACCATAA CTACAATACT CTAAAAGGTA GTCTTGTAAG CGACTGATTT 4620
 TTTCAAGTTT TTCTTCTGTT CGGTCATCCT TGTAAAAACC TCGTAGGCGA AGCTGTTCTG 4680
 TGCTCCAGTC CCCCACGATA TAATCAAAC TGGTTAATAC TTCTGAAAA CGCTGATTAA 4740
 AAGTCGTAC ATCAAAGGCA TCCTTGATAT TTTCAACCA GAAAAAGCT ATCCCTTCCG 4800
 TTTCGACCTT GTCCCGTGT AAATGGAAC CCGGACCAGG AACTTGTTA TAGTTGTATA 4860
 ATTCAGGTGC AATTCTTTT CGCATAGATA TCCTTTTTC ACGATTACTT AATACTTTAT 4920
 TCTACCATAA TTTCTAGCAG TTAGCACGTT TCTCATAAAA ATGAAAAAG TCTGACGATT 4980
 TTGTCAGACC AGAATCTTAT AACCTAAAA GAGAAGAACA ATTCTCCCT CCAACTATCA 5040
 TTATTTAGCA GCTGCGTACA ATTCATCTAC TTTATTCCAG TTGATTACTG AAAAGAAAGC 5100
 TTTGATGTAG TCAGGACGCA CGTTGCGGTA TTTCACGTAG TAAGCATGTT CCCAAACGTC 5160
 CAAGCCCAAG ATTGGTTTTT TACCTTCTGA GATTGGGTG TCTTGGTTTG CTGTTGAAGT 5220
 CACTTCAAGT TTCCCTTCTT TGTGACAAC CAACCATGCC CAACCTGAAC CAAAACGAGT 5280
 TGTGCTGCT GCAGTGAAG CTGCTTGGA TTCTCAAAT GAACCAAATG TTGCATCGAT 5340
 TGCTGCTGCC AGTTCTGCTG AAGGAGCTGT TTTCTCGGA GTCATCAATT CCCAGAAAAG 5400
 AGCGTGGTTC AAGTGTCGC CACCATTGTT GATAAGTGCT TGACGGATAT CAGCTGGGAT 5460
 AGATTCTACA TCAGCAAGCA AGGCTTCAAG GTCTTCACCG ATTTCAAGGT GTTTTCTAA 5520
 AGCTGCATTG GCATTGTTGA CATAAGTTTG ATGGTGTT 5558

(2) INFORMATION FOR SEQ ID NO: 104:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 6735 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 104:

GGAATTGTAA ATATCATATT GTTTTGCAC CCAAATATCG TCGTCAAATC ATTTATGGCA 60
 GATACAAAGC TAGTATCGGA AGAATCATAC GTGACTTATG TGAGCGTAAG GGTGTAATAA 120
 TCCATGAAGC GAATGCTTGT TCAGACCATA TTCACATGCT TATCAGTATT CCTCCGAAAC 180

| | | |
|------------|-------------|------------|
| 774 | | |
| TTAGTGTTTC | GTCCCTTATG | GGCTATTTAA |
| AGCATGCGAA | TTTAAAATAC | AAATATGGCA |
| TAGATACGGT | AGGCCGTAAT | CAGAAAGTGA |
| AAGACAGAGT | AGCAGACCAG | CTCACGTTAT |
| TAAATAAGAG | GAAGTAACTA | AGGTGCTTTA |
| AGCTATTTTC | GTGGGCCTTT | GGCCCTGGCC |
| CCACCAGTTC | ACACTGGTGG | TTTTGATTTA |
| TATAAAGGAT | GGTAAAATTC | CTGTTGTCCG |
| TATGGTCTAT | ACTTTTCTTT | AGGAGAAAGC |
| GAGGATCATG | ATCTGCCCCA | AAAGCAAATA |
| TATGCCAAAA | TTGAACGGGG | TGAGCATGCG |
| GATTTCTATG | ACGTCAGTAC | AGACTATTTA |
| CGCTTTAGAA | AATAATCTCC | TCAATTTTCT |
| TGCCCTTTGA | CAACTGAATA | GCCTAAAATG |
| ATGGCTCGCC | ATGATAAGAG | CGATTTTAAA |
| TGCCATGATA | CAAAATGATAT | ACAATGATAC |
| AGCAGCAAGT | GAAATCTTTA | TGATGACTTC |
| TGTTAGATAA | ACGCAATTAA | TCCTCAAAAG |
| CATCACGTGG | AGTGTGTAAG | CTGTGTGCTA |
| AATAGACTTT | CTGCGAAACA | AAAATATAAT |
| AACAATTGAG | CGATAGCCGT | TTCAAGATCC |
| AGATGTTAGC | TGTGTTAAAA | ACAGCTTATC |
| GCAAATTAAG | CCTAGACGAT | CTCCTTATGG |
| CTTATGAACA | AATTGCGGCT | GATTTTGGCA |
| AATGGGTGA | AGCAACTCTT | ATTCAAAATG |
| TGTAAAAACA | GTAAAATTCG | AAGGATTGTA |
| GGTATAATAG | CAATCAAAAC | TAGAAAATAA |
| CTAGTAGAGT | GGTGATACTA | TGAAGATTAG |
| GATTCCCTAC | TTGCTTTTAT | CTATTTTGGG |
| TATTTTAATT | GAAGAAGGCA | AGAGCGCCTT |

240

300

360

420

480

540

600

660

720

780

840

900

960

1020

1080

1140

1200

1260

1320

1380

1440

1500

1560

1620

1680

1740

1800

1860

1920

1980

775

| | |
|---|------|
| GATTGTTAGT TTGATACTGA TTGCCTTAAT TTATAAATTG AGACTAGATT TTTTGAGAAA | 2040 |
| TGAGCGACTA ATCATTTTAG TTATATTAAT AGAAATGCTT TTATTGTTCT TGGCTCGTTT | 2100 |
| TATTGGTATT TCCGTAAACG GGGCATACGG TTGGATTTCG GTTGCAGGAA TAACTATTCA | 2160 |
| GCCAGCTGAG TACTTAAAAA TCATTATTAT TTGGTATTTA GCTCACCGAT TCTCCAAACA | 2220 |
| GCAAGAAGAA ATAGCTACTT ATGATTTTCA AGTTTGTACT CAAAATCAAT GGCTTCCCCG | 2280 |
| TGCTTTTAAT GATTGGCGAT TCGTTCTCCT AGTTCTGATT GGAAGTTTGG GAATTTTCCC | 2340 |
| TGATTTAGGA AATGCGACTA TTTTAGTCTT GGTTCCTTG ATTATGTATA CAGTTAGTGG | 2400 |
| AATCGCTTAT CGCTGGTTT CAACCATCT GCGCTCGTA TCTGCCGCTT CTGTCTTGT | 2460 |
| CTTGACCACT ATCAGCCTAA TCGGTGTGA GACCTTTTCA AAAATTCAG TATTCGGCTA | 2520 |
| TGTAGCCAAG CGCTTAGTG CCTTTTAA TCCTTTGCC GATCGTCTG ATGCAGGTCA | 2580 |
| CCAGTTAGCT AATTCTTATT TTGCCATGGT CAATGGCGGT TGGTTTGGTC TAGGTCTTGG | 2640 |
| AAACTCGATT GAAAAACGAG GTTATTTGCC AGAAGCTCAT ACAGACTTTG TCTTTTCTAT | 2700 |
| CGTGATTGAA GAATTTGGCT TTGTGGTGC CAGTCTTATT TTAGCTCTCT TGTTTTTCAT | 2760 |
| GATTTTGGCG ATTATCTTGG TCGGTATCCG AGCGGAGAAT CCTTCAATG CCATGGTTGC | 2820 |
| ACTCGGTGTC GGAGGGATGA TGTGGTTCA GGTATTGTC AATATCGGAG GGATTTCGGG | 2880 |
| CTTGATTCCA TCTACAGGAG TGACTTTCCC CTCTTATCC CAGGGTGGAA ATAGTCTTCT | 2940 |
| AGTCTTATCA GTGGCAGTAG CCTTTGTCTT AAATATTGAT GCCAGTGAAA AACGCGCTAA | 3000 |
| ATTGTACCGA GAATTGAAA ATCAACCAAT GAACCTTCTG TTGAAGTAGG ATAAAGAAAG | 3060 |
| GATAGTTTAT GTCTCTTCAA AAATTAGAAA ATTATAGTAA TAAAAGTGT GTGCAAGAAG | 3120 |
| AAGTCTTGAT TCTAACAGAA TTACTGGAAG ATATTACTAA AAATATGCTT GCCCCAGAGA | 3180 |
| CCTTTGAAA AATAATACAG TTGAAAGAAT TATCAACGCA GGAAGATTAT CAAGTCTAA | 3240 |
| ACCGTCTAGT GACTAGCTTA TCAATGATG AAATGGTCTA TATTTACGC TATTTCTCTA | 3300 |
| TCTTGCTCT TTTGATTAAT ATTTAGAGG ATGTGGATT AGCTTATGAA ATCAATCATC | 3360 |
| AAAATAATAT TGATCAGGAC TATTTAGGTA AATTATCTAC AACGATTAAA TTGGTAGCAG | 3420 |
| AAAAGGAAAA TGCCGTTGAG ATCCTAGAAC ACTTGAATGT TGTCCCTGTT TTGACAGCCC | 3480 |
| ATCCAACACA AGTGCAACGC AAAAGTATGT TGGATTTAAC AAATCATATT CATAGTCTTT | 3540 |
| TGCGTAAATA CCGTGATGTT AAGTTGGGT TGATCAATA AGATAAATGG TACAATGATT | 3600 |
| TGCGTCGTTA CATCGAAATT ATCATGCAGA CAGACATGAT TCGTGAGAAA AAATTAAAAG | 3660 |
| TGACTAACGA AATCACGAAT GCTATGGAAT ATTATAACAG CTCCTTTTGT AAAGCTGTAC | 3720 |

776
CTCATTGAC GACGAGTAT AAGCGCTTAG CGCAAGCGCA TGGTCTGAAT TTA AACACAGG 3780
CTAAACCAAT CACCATGGGT ATGTGGATAG GTGGTGACCG TGATGGAAAT CCATTTGTTA 3840
CAGCAAAGAC CTTGAAGCAG TCTGCACTCA CTCAGTGTGA AGTCATCATG AACTACTATG 3900
ATAAAAAGAT TTACCAACTT TATCGTGAAT TTTCTCTTTC AACTAGCATT GTCAACGTCA 3960
GCAAGCAAGT CAGAGAAATG GCTCGTCAAT CCAAGGATAA CTCGATTTAC CGCGAAAAAG 4020
AGCTTTACCG TCGTGCCTTG TTTGATATTC AATCAAAAAT TCAGGCAACT AAAACCTATC 4080
TGATTGAGGA TGAAGAAGTT GGGACTCGTT ATGAAACCGC CAATGATTTC TACAAGGATT 4140
TGATTGCCAT TCGAGATTCT CTACTAGAAA ATAAGGGCGA GTCCTTGATT TCAGGTGATT 4200
TTGTGGAATT ATTGCAGGCA GTAGAGATAT TTGGTTTTTA CTTAGCATCA ATTGATATGC 4260
GACAAGACTC TAGCGTCTAT GAAGCCTGTG TGGCAGAACT CTTGAAATCA GCAGGAATTC 4320
ATTCTCGTTA TAGCGAGTTG AGCGAAGAAG AAAAGTGTGA CCTTCTCTTG AAAGAATTAG 4380
AAGAAGATCC CCGAATTCTT TCTGCGACTC ACGCAGAAAA ATCAGAATTA TTAGCAAAAAG 4440
AATTAGCTAT TTTAAGACG GCTCGTGTTC TGAAAGATAA GTTGGGAGAT GATGTCATCC 4500
GTCAGACCAT CATTTACAT GCAACCAGCC TTTCTGATAT GCTAGAATTA GCTATTCTGT 4560
TAAAGAAGT AGGACTGGTG GATACGGAAA GGGCGCGTGT TCAGATTGTT CCCCTTTTGT 4620
AAACAATTGA AGACTTGAT CATTCAGAGG AAACAATGAG AAAATATCTT TCTCTTAGCC 4680
TTGCCAAAAA ATGGATTGAC TCACGAAATA ACTACCAAGA AATCATGCTT GGCTACTCTG 4740
ACAGTAATAA AGATGGCGGT TACTTGTCAT CATGTTGGAC CCTCTACAAG GCTCAACAAC 4800
AATTGACTGC TATTGGAGAT GAATTTGGCG TTAAGGTTAC CTTCTCCAT GGTCTGCTG 4860
GTACTGTCGG TCGTGGTGGT GGGCCAACCT ATGAAGCCAT TACATCTCAA CCGCTCAAGT 4920
CTATCAAGGA TCGTATCCGC TTGACGGAGC AGGTGAAGT AATTGGGAAT AAATACGTA 4980
ACAAAGACGC CGCTTACTAT AACCTTGAAA TGCTAGTATC GGCAGCTATT AACCGTATGA 5040
TTACTCAGAA GAAGAGCGAT ACCAATACCC CAAATCGTTA TGAAACCATT ATGGATCAAG 5100
TAGTGGACCG TAGTTACGAT ATCTACCGTG ATTTGGTCTT TGGTAATGAG CATTTCTATG 5160
ATTATTTCTT CGAGTCAAGT CCAATCAAGG CTATTTCAAG TTTTAATATT GGTCTCTGTC 5220
CAGCCGCTCG TAAGACTATT ACTGAAATCG GTGGTTTTCG TGCCATCCCT TGGGTATTCT 5280
CATGGTCACA GAGTCGTGTT ATGTTCCCTG GATGGTACGG GGTGGTTCA AGCTTCAAGG 5340
AATTTATCAA TAAAAATCCA GAGAATATTG CTATCTTACG AGATATGTAC CAAAATTGGC 5400
CTTCTTCCA ATCGCTTCTT TCAAATGTTG ATATGGTTTT GTCAAAAATCA AATATGAATA 5460
TTGCTTTTGA ATATGCTAAA CTTTGTGAAG ACGAGCAAGT TAAGGCCATC TATGAGACTA 5520

777

TTTTAAATGA ATGGCAAGTT ACTAAGAACG TTATCTTGGC TATTGAAGGA CATGACGAAC 5580
 TCTTAGCTGA CAATCCATAT CTAAGAGCTA GTCTGGATTA CCGTATGCCT TACTTTAATA 5640
 TTCTCAACTA TATTCAGTTG GAGTTGATTA AACGCCAACG TCGTGGAGAA TTGTCCAGTG 5700
 ATCAAGAACG ATTGATTCAT ATCACCATCA ACGGAATTGC GACAGGATTG CGTAATTCAG 5760
 GTTGATAATT TTCAAGAGTG AATGCTAAAA GTGAATATCA AAAAAATTCT AATAGACTAT 5820
 TGACAAGTAG TTTAAAAATG ATATAATTTA ACCATTCAGA AAAGTAATCA TACAAACTTT 5880
 TTAGAGAGTC TGTGGTAGCT GAAACAGAT AAGTGGCAAT GATGAAAATT GGGCTGAATG 5940
 CTATTTAGAA TTTGAAATTA TAAAAATTCG GTAAGCACAC CTTACAGTGC ATCTCGTTAT 6000
 TGCAGACTG AGCGATAGGG AAATTCCTTA TAATTGAGGT GGTACCGCGC ATCGACGTCC 6060
 TCACACAAGT TTTTGTGTG AGGATTTTTT TGATGGAGGT TAGTATGGAA AGAAAACGAT 6120
 GGCGTCGCTT GTTTAGATAA GTGAAATATG TTAAAGGAAA TAAAAAGGAG AAACAGAATG 6180
 AAAAAATAAC GTTTAATTGG AATTATTGCT GCATTAGCAG TCTTAGTAGC AGGAAGCTTG 6240
 ATTTATTCTT CAATGAATAA ATCAGAAGCT CAGAATAATA AGGATGAGAA GAAAATAACC 6300
 AAGATTGGTG TGCTTCAATT TGTGAGCCAT CCATCCCTTG ATTTGATTTA TAAAGGGATC 6360
 CAAGATGGAC TTGCAGAAGA AGGATATAAA GATGATCAAG TTTAAATTGA TTTTATGAAC 6420
 TCAGAAGGTG ACCAAAGTAA GGTGCGACA ATGAGTAAAC AATTGGTTGC AAATGGGAAT 6480
 GACCTTGTGG TTGGTATCGC AACACCAGCA GCCCAAGGGT TGGCTAGTGC AACAAAAGAC 6540
 CTACCGGTTA TCATGGCCGC TATTACAGAC CCAATTGGTG CTAAGTTGGT TAAAGATTG 6600
 AAAAAACCAG GTGGCAACGT TACAGGGGTA TCTGACCACA ATCCAGCTCA ACAACAAGTT 6660
 GAAGTCATCA AGGCTCTGAC ACCGAATGTG AAAACAATCG GAGCTCTTTA CTCAAGTAGC 6720
 GAAGACAATT CAAA 6735

(2) INFORMATION FOR SEQ ID NO: 105:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 6516 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 105:

CTAGAGGATC CCAGCAGGTA AATTGGCTTC AGCTGGCAAA AAAGTTGCCC TCGTTGAACG 60
 CAGCAAGGCT ATGTACGGTG GAACTTGAT CAACATTGGT TGTATCCCAA CTAACCTT 120

778

| | |
|--|------|
| GCTAGTTGCT GCTGAAAAGG ACTTGTCTTT TGAAGAAGTC ATTGCTACTA AAAACACGAT | 180 |
| CACTGGTCGC CTCAACGGTA AAAACTATGC GACTGTTGCT GGTACAGGCG TAGATATCTT | 240 |
| TGATGCGGAA GCTCACTTCC TTTCAAATAA AGTCATCGAA ATCCAAGCTG GTGATGAAAA | 300 |
| GAAAGAACTG ACTGCTGAAA CAATCGTCAT CAACACTGGT GCTGTTTCAA ACGTCTTGCC | 360 |
| AATCCCTGGA CTTGCTACAA GCAAAAACAT CTTTGACTCA ACAGGTATCC AAAGCTTGGA | 420 |
| CAAATTACCT GAAAAACTTG GAATCCTTGG TGGCGGAAAT ATCGGTCTTG AATTGCGCGG | 480 |
| CCTTTACAAC AAACCTGGAA GCAAGGTCAC AGTCCTAGAT GCCTTGGATA CATTCCTACC | 540 |
| TCGTGCAGAA CCTTCCATCG CAGCTCTTGC TAAACAATAC ATGGAAGAAG ATGGCATTTGA | 600 |
| ATTGCTTCAA AATATCCATA CTAAGTAAAT CAAAAACGAT GGTGACCAAG TGCTTGTCTG | 660 |
| AACTGAAGAC GAAACTTACC GTTTCGACGC CCTTCTCTAC GCAACTGGAC GCAAACCAAA | 720 |
| TGTAGAACCA CTTCAACTTG AAAATACAGA TATTGAACTA ACTGAACGTG GTGCTATTAA | 780 |
| AGTAGACAAA CACTGTCAAA CAAACGTTCC TGGTGTCTTT GCAGTTGGAG ATGTCAACGG | 840 |
| TGGCCTTCAA TTTACTTACA TTCACTTGA TGACTTCCGT GTTGTTTACA GCTACCTTGC | 900 |
| TGGAGATGGC AGCTATACAC TTGAAGACCG TCTCAATGTG CCAAATACTA TGTTTCATCAC | 960 |
| ACCTGCACTT TCACAAGTTG GTTTGACTGA AAGCCAAGCA GCTGATTTGA AACTTCCATA | 1020 |
| CGCTGTAAAG GAAATCCCGG TTGCAGCAAT GCCTCGTGGT CACGTAAATG GAGACCTTCG | 1080 |
| CGGTGCCTTC AAAGCTGTTG TCAATACTGA AACAAAAGAA ATTCTTGGAG CAAGCATCTT | 1140 |
| CTCAGAAGGT TCTCAAGAAA TCATCAACAT CATCACTGTT GCTATGGACA ACAAGATTCC | 1200 |
| TTACACTTAC TTCACAAAC AAATCTTCAC TCACCCAACC TTGGCTGAGA ACTTGAATGA | 1260 |
| CTTGTTTGCG ATTTAAGTTG AGATTTAATC GTATCGAACA GCCCTCTTTG GGCTGTTTTT | 1320 |
| ACTTCTGCGG AATCTCAAAT CTGTCTTTCT CCTCTTTTAT GATATAATAG AAACATGAAC | 1380 |
| TTAAAACTA CTTTGGGCCT TCTTGCTGGG CGTTCTTCCC ACTTCGTTTT AAGCCGTCTT | 1440 |
| GGACGTGGAA GTACGCTCCC AGGGAAAGTC GCCCTTCAAT TTGATAAAGA TATTTTACAA | 1500 |
| AACCTAGCTA AGAACTACGA GATTGTCGTT GTCACTGGAA CAAATGGAAA AACCTGACA | 1560 |
| ACTGCCCTCA CTGTGGGCAT TTAAAAAGAG GTTTATGGTC AAGTTCTAAC CAACCCAAGC | 1620 |
| GGTGCCAACA TGATTACAGG GATTGCAACA ACCTTCCTAA CAGCCAAATC TTCTAAAACT | 1680 |
| GGGAAAAATA TTGCCGTCCT CGAAATTGAC GAAGCCAGTC TATCTCGTAT CTGTGACTAT | 1740 |
| ATCCAGCCTA GTCTTTTGT CATTACTAAT ATCTTCCGTG ACCAGATGGA CCGTTTCGGT | 1800 |
| GAAATCTATA CTACCTATAA CATGATATTG GATGCCATTC GGAAAGTTCC AACTGCTACT | 1860 |
| GTTCTCCTTA ACGGAGACAG TCCACTTTTC TACAAGCCAA CTATTCCAAA CCCTATAGAG | 1920 |

779

| | |
|--|------|
| TATTTTGGTT TTGACTTGGA AAAGGGACCA GCCCAACTGG CTCACTACAA TACCGAAGGG | 1980 |
| ATTCTCTGTC CTGACTGCCA AGGCATCCTC AAATATGAGC ATAATACCTA TGCAAACTTG | 2040 |
| GGTGCCTATA TCTGTGAAGG TTGTGGATGT AAACGTCCTG ATCTCGACTA TCGTTTGACA | 2100 |
| AAACTGGTTG AGTTGACCAA CAATCGCTCT CGCTTTGTCA TAGACGGCCA AGAATACGGT | 2160 |
| ATCCAAATCG GCGGGCTCTA TAATATCTAT AACGCCCTAG CTGCTGTGGC CATCGCCCGT | 2220 |
| TTCTTAGGTG CCGATTGCGA ACTCATCAA CAGGGATTG ACAAGAGCCG TGCTGTCTTT | 2280 |
| GGACGCCAAG AAACCTTTCA TATCGGTGAC AAGGAATGTA CCCTTGTCTT GATTAAAAAT | 2340 |
| CCAGTCGGTG CAACCCAAGC TATCGAAATG ATCAAACCTAG CACCTTATCC ATTTAGCCTA | 2400 |
| TCTGTCTCTC TTAATGCCAA CTATGCAGAT GGAATTGACA CTAGCTGGAT CTGGGATGCA | 2460 |
| GACTTTGAAC AAATCACTGA CATGGACATT CCTGAAATCA ACGCTGGCGG TGTTCGTCAT | 2520 |
| TCTGAAATCG CTCGTCGCCT CCGAGTGACT GGCTATCCAG CTGAGAAAAT CACTGAAACG | 2580 |
| AGTAATCTGG AGCAAGTTCT CAAGACCATT GAGAATCAAG ACTGCAAGCA TGCCTATATT | 2640 |
| CTGGCAACTT ATACTGCCAT GCTGGAATTT CGTGAAGTGC TGGCTAGTCG TCAGATTGTT | 2700 |
| AGAAAGGAGA TGAACATAAG GTTTATACTT CACTTTCCTC AAAAGATGGC AATTACCCCT | 2760 |
| ATCAGCTCAA CATTGCCAC CTCTACGGAA ATCTCATGAA TACTACGGGG ACAATGGAAA | 2820 |
| CATCCTCATG CTCAAGTATG TGGCTGAAAA ACTGGGAGCC CATGTGACCG TTGACATCGT | 2880 |
| TTCTCTCCAT GATGACTTTG ATGAAAATCA CTACGACATC GCCTTTTTCG GTGGTGGTCA | 2940 |
| AGACTTTGAA CAAAGTATCA TTGCAGACGA CCTACCTGCT AAAAAAGAGA GCATTGACAA | 3000 |
| CTACATCCAA AACGACGGTG TAGTTCTGGC TATCTGCGGT GGTTCCTAAC TATTGGGTCA | 3060 |
| ATATTATGTT GAAGCTTCAG GAAAACGTAT CGAAGGGCTA GGGGTCATGG GACACTACAC | 3120 |
| GCTCAACCAG ACCAATAACC GTTTTATCGG TGACATCAAG ATTCACAATG AAGATTTGCA | 3180 |
| TGAAACCTAC TATGGATTG AAAATCACCA AGGTCGTACC TTCCTCTCTG ATGACCAAAA | 3240 |
| ACCGCTGGGA CAGGTGTCT ATGGAAATGG AAACAACGAA GAAAAGGTCG GTGAAGGGGT | 3300 |
| TCATTATAAG AATGTCTTTG GTTCCTACTT CCACGGGCCT ATCCTCTCTC GTAATGCCAA | 3360 |
| TCITGGCTTAT CGCCTAGTTA CTACTGCCCT CAAGAAGAAA TATGGTCAGG ACATCCAAC | 3420 |
| CCCTGCCTAT GAGGACATTC TCAGCCAAGA AATCGCTGAA GAGTACAGTG ACGTCAAAA | 3480 |
| CAAGGCTGAC TTTTCTTAAA CAAAGGAAAA TGATATCAAA GAACTCCGTT ATCTTGTCGG | 3540 |
| AGTTTTTTGT CTTTTCTTTT ACCCTTCTCC CTTGCATTTT CTCTCATTTT TTGCCAAAAT | 3600 |
| AGAGGGGTAG AAAGAAGGTA GCATATGTCT AAATTACAAC AAATCCTAAC ATATCTTGAA | 3660 |

| | | | | | |
|-------------|------------|-------------|------------|------------|------------------|
| 780 | | | | | |
| TCAGAAAAAC | TAGACGTCGC | TGTCGTATCT | GACCCCGTCA | CAATCAATTA | CCTCACTGGT 3720 |
| TTTTACAGTG | ATCCCCATGA | ACGCCAAATG | TTCTCTTTTG | TCCTAGCAGA | TCAGGAACCT 3780 |
| CTCCTCTTTG | TCCCAGCTCT | TGAAGTAGAA | CGTGCAAGTA | GCACCGTTTC | CTTCCCAGTA 3840 |
| GTGGGGCTATG | TCGATTCTGA | AAATCCATGG | CAAAAAATCA | AACATGCTCT | TCCACAACCT 3900 |
| GACTTCAAAC | GTGTCGCTGT | TGAGTTTGAC | AATCTCATCT | TGACCAAATA | CCATGGTTTG 3960 |
| AAAACAGTTT | TTGAGACTGC | TGAGTTTGAC | AACCTCACTC | CTCGTATCCA | ACGCATGCGC 4020 |
| CTCATCAAAT | CAGCTGATGA | AGTGCAAAAA | ATGATGGTTG | CAGGTCTTTA | TGCTGACAAG 4080 |
| GCTGTTTCATG | TTGGTTTGA | CAATATTTCT | CTTGATAAGA | CTGAGACAGA | TATCATCGCA 4140 |
| CAAATCGACT | TTGCCATGAA | ACGTGAAGGT | TATGAAATGA | GCTTTGATAC | CATGGTCTTG 4200 |
| ACTGGTGATA | ATGCTGCGAA | TCCACACGGC | ATTCCAGCAG | CTAATAAGGT | TGAAAATGAT 4260 |
| GCTCTTCTCC | TCTTTGACCT | GGGTGTTCTG | GTCAATGGCT | ATGCGTCAGA | TATGACTCGT 4320 |
| ACAGTCGCTG | TCGGCAAACC | AGACCAATTC | AAGAAAGATA | TTTACAACCT | GACTCTTGAA 4380 |
| GCCCAACAAG | CTGCTCTTGA | CTTTATCAAG | CCAGGTGTGA | CTGCTCATGA | AGTGGACCGC 4440 |
| GCTGCCCGTG | AGGTCATCGA | AAAAGCTGGT | TATGGTGAGT | ACTTCAACCA | CCGTCTCGGG 4500 |
| CATGGTATCG | GTATGGATGT | CCATGAATTC | CCATCTATCA | TGGAAGGAAA | CGACATGGTC 4560 |
| ATCGAAGAAG | GCATGTGCTT | CTCTGTTGAA | CCAGGTATCT | ATATCCCTGG | TAAAGTCGGT 4620 |
| GTTCGTATTG | AAGACTGCGG | TGTTGTTACC | AAGGATGGCT | TCAACCTCTT | TACAAGCACC 4680 |
| AGCAAAGATT | TGCTTTATTT | TGATTAAACT | ATATAGCCCC | TATGCTTTCC | TTTCAAAATA 4740 |
| TCTAGGGGCT | ATTTTATGT | CATTTTCTG | CTATTATGCT | AAAGAAATTG | GCTGCAATAA 4800 |
| TCTAACCCTA | AGTGTCTGGA | ATGATAACGA | GGGTGCTCTC | CGCTTTTATC | AAAGACAAGG 4860 |
| GATGAAACCC | CAAGAAACAA | CAATGGAAAT | GATAATTGAT | TAAGAAGTCA | TCTATCAAAA 4920 |
| GATGTTAGAA | AAAGTTCAAT | TTCAC TAGAA | AATGAGGAAA | ATCTCCCCAC | AATAAAAACGC 4980 |
| ATAGTATCAG | GTATTGTGTA | CTGACCCCAA | ACAGTTAGAC | AATTAATTTA | TCCGAAGGAT 5040 |
| TTAGTTCTGT | ACTGCACAGG | ACTAAGTCCT | TTTAGTTTTA | CCTTAATTCG | TTTGTGTGTG 5100 |
| TAGTAATCAA | TATAGTCTAT | AATGACTTGT | TCCAATTGGT | TAAGTGATTT | AAATGTTTTC 5160 |
| TCATAGCCAT | AAAACATTTC | GGATTTTAAA | ATGCCAAAGA | AAGATTCCAT | CATACCGTTG 5220 |
| TCTTGGCTGT | TTCCCTTGCG | TGACATAGAT | GCTTGAATTC | CCTTATTCTC | TAGGAACCGA 5280 |
| TGATAAGAAT | CGTGTGGTA | TTGCCAGCCT | TGGTCACTAT | GGAGAATCGT | ATTCTCGTAG 5340 |
| TGCTTCTCTT | TGAATGCCTG | TTCCAACATT | GTTTGTAATT | ATTCTAAATT | AGGCGAACAA 5400 |
| GAAAGATTAA | AAGCAATAAT | TTCGCTGTTA | AAGCCATCTA | AAACTGGTGA | TAAGTAAAGC 5460 |

781

TTTTGAGTAC TTGCTGGAAT GGCAAATTCA GTCACATCTG TGTAGCACTT TTCCATTGTT 5520
 TTAGAGCCTT CAAATTGGGC TTGAATGAGA TTCTCTGCCT TCTTACCAAC GTCTCCCTTA 5580
 TGAGAAGAAT ATTTTCGTTT CTTTCGCATT TTAGCTTGTA AATTGAGTAC TTTCATCAAG 5640
 CCTTGAACTC TTTTATGATT TACCAGATAA CCACGATTTT TTAGTTCTAA ATGAACCCGG 5700
 CGATAAGCAT AATTTCCCTT GTGTTGCGATA AAGATGGATT GAATTTCACT TTTAAGCTCT 5760
 TGGTCTTTAT CTGTTTGTG TAGCTGTTTC AAGTGATAGT AGTAGGTCCA ACGAGCTAGT 5820
 TTAATGGCTT CTAGAAGAAG ATCTAACGAA AACTCAGTCA TTAATTCCTG AACAAATTCT 5880
 GTCTTTCTTC TTTCTCTTTT TCCTCCTTCA ATCGGAGTTC TCTTAACCTT TTTAGGATGG 5940
 CATTTCTCCG TCTCAGGTAC TCTCCCTCTT GTTTTCTCAA CAATAGTATA CCCGTTTTTC 6000
 CTGTATTGTG CTAGCCAGTT AAGAAGTATC GTACGACTTG GGAGACCGTA TTCAAGAGAA 6060
 ACTCTATCTT TAGTCCAGCC TTCATGTCAG ACTTTATTAA CCCCAATTAT TCACCCCAA 6120
 TCTAAAAACC ATCCAGAATC CTTGCCTTAG CTTAGATCCT GGATGGTTTC TTTTTCACC 6180
 CAATGGGTGT TTTTACTAG AAAAAAAGA GTTTCCCTT TATGGTATAA GTGTAGAAAA 6240
 AAACACAAAA AGAAAGGAAA CTCACATGAA CAGTTTACCA AATCATCACT TCCAAAACAA 6300
 GTCTTTTAC CAACTATCTT TCGATGGAGG TCATTTAACC CAGTATGGTG GTCTTATCTT 6360
 TTTTCAGGAA CTTTTTCCC AGTTGAACT AAAAGAGCGG ATTTCTAAGT ATTTAGTAAC 6420
 GAATGAmCAA CGCCGCTACT GTCGTTATTC GGATTGAGAT ATCCWTGTCC AGTTCCTCTT 6480
 TCAACTGTTA ACAGGTTATG GAACGGAATA TGCTTG 6516

(2) INFORMATION FOR SEQ ID NO: 106:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 14654 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 106:

TTTTCAACCC ATATCGTGGC TCCTGAATAC TACTTACTGA CAACTATGCT ATCAGAGACT 60
 TCTCTACTTG TTTTCTATAT CATTTTCATC CATAGAAAAC AACTCATCCA CTTGGGACAT 120
 ATCTTTAGCT ATACTGTTTCG ATACTCTCTC TTTTCACTT CCTTTGTAGC AATTATTTTC 180
 CTGATTAATT TCGTGTATCC TGTAATATG GTCATTAATT TGCCATTTT GATTAATACT 240
 GGTTTGATTG TCTTGCTATC AGCTATCTCT TATATTAGTC TACTTGTCTT CACAAAAGAT 300

782

| | |
|--|------|
| AGCATTTTCT ATGAATTTT AAACCATGTC CTAGCCTTAA AAAATAAATT TAAAAAATCA | 360 |
| TAGGAGTTTA AAATGAAACA ACTAACCGTT GAAGATGCCA AACAAATTGA ATTAGAAATT | 420 |
| TTGGATTATA TTGATACTCT CTGTAAAAAG CACAATATCA ACTATATTAT TAACTACGGT | 480 |
| ACTCTGATTG GGGCGGTTCG ACATGAGGGC TTTATCCCTT GGGACGACGA TATTGATCTG | 540 |
| TCCATGCCTA GAGAAGACTA CCAACGATTT ATTAACATTT TTCAAAGGA AAAAAGCAAG | 600 |
| TATAAGCTCC TATCCTTAGA AACTGATAAG AACTACTTTA ACAACTTTAT CAAGATAACC | 660 |
| GACAGTACGA CTAAAATTAT TGATACTCGA AATACAAAA CCTATGAGTC TGGTATCTTT | 720 |
| ATCGATATTT TCCCTATAGA TCGCTTTGAT GATCCTAAGG TCATTGATAC TTGTTATAAA | 780 |
| CTGGAAAGCT TCAAACTGCT GTCTTTCAGT AAACATAAAA ATATTGTCTA TAAGGATAGC | 840 |
| CTTTTAAAAG ATTGGATACG AACAGCCTTC TGGTTACTCC TTCGACCGGT TTCTCCTCGT | 900 |
| TATTTTGCAA ATAAAATCGA GAAAGAAATT CAAAAATATA GTCGTGAAAA TGGGCAATAT | 960 |
| ATGGCTTTTA TCCCTTCAAA ATTTAAGGAA AAGGAAGTCT TCCCAAGTGG TACCTTTGAT | 1020 |
| AAAACAATCG ATTTACCCTT TGAGAATTTA AGCCTTCCTG CACCTGAAAA ATTTGATACT | 1080 |
| ATTTTGACAC AATTTTATGG AGATTATATG ACCCTACCAC CAGAAGAAAA ACGCTTCTAC | 1140 |
| ACTCATGAAT TTCACGCTTA TAAATTGGAG GATTAGGATG CAATATTTAG AAAAAAAGA | 1200 |
| AATTAAAGAA ATTCAACTAG CCCTGCTGGA CTATATTGAT GAGACTTGTA AGAAACATGA | 1260 |
| TATTCCTTAT TTTCTCAGTT ATGGAACCAT GCTTGGAGCC ATCCGCCACA AAGGTATGAT | 1320 |
| TCCTTGGGAT GATGATATTG ATATTTCCCT TTATCGTGAG GATTATGAGC GTTACTGAA | 1380 |
| GATTATTGAA GAAGAAAATC ACCCTCGCTA CAAGGTTCTT TCCTACGATA CATCTTCTTG | 1440 |
| GTACTTCCAT AATTTTCGCAT CGATTTTGGA CACTTCTACT GTTATAGAAG ACCATGTTAA | 1500 |
| GTACAAGCGT CATGATACCA GCCTTTTCAT CGATGTCTT CCAATTGATC GATTTACAGA | 1560 |
| CTTGAGCATT GTCGACAAGA GCTATAAGTA TGTGGCTCTT CGTCAACTAG CTTATATCAA | 1620 |
| AAAATCACGA GCAGTTCACG GTGATAGCAA ACTAAAAGAT TTTCTTAGAT TATGTAGCTG | 1680 |
| GTACGCTCTC CGATTTGTCA ATCCTCGCTA CTTTACAAG AAAATTGATC AACTAGTCAA | 1740 |
| AAATGCTGTA ACCAACACTC CTCAATATGA AGGAGGAGTT GGGATCGGTA AGGAAGGGAT | 1800 |
| GAAAGAAATC TTCCCAAGTT ATACCTTTAA AGAACTGATT TTAAGTGAAG TTGAGGGCCG | 1860 |
| TATGTTGCCT GTTCCCAAAA AATATGACCA ATTTTAAACC CAGATGTATG GCGATTATAT | 1920 |
| GACACCACCA TCAAAAGAAA TGCAAGAGTG GTATAGTCAT AGCATTAAG CTTATCGCAA | 1980 |
| AAACTGATTG AGGGGGATTA TACAACTAC TAAGATAGAG GTTATTCAAA AACATAATTT | 2040 |
| TAGTAGAAAA TGAAATACAT ATTCCACAA TAAACGCAT CATATCAAGG TTTTGA AAAA | 2100 |

783

| | |
|--|------|
| ACCTTGATAT GATGCGTTTT ATAATTTTAA AGACTTTTTT CTATAGTAGA TTGAAATAAG | 2160 |
| ATGCGAACAA ATCAATTAGA AAATTCAAAT TAATTTATAG AAATATTTTA GTATTCCTGT | 2220 |
| GTA CTGTTCT AAATTCAGTC TGCTATATCT TATTTTCTA TTAAATCGC TTCTGTAACA | 2280 |
| AAGCTACGAC TTCAAGTAC CTTAAGCATG GCATTAGCTG TATCTAGCGC TGTGAAGAGG | 2340 |
| GGCACCCCGT GTTCAATGGC TGAACGACGA ATTTGCTCAC CATCTTCGTC AGCAGTTCGT | 2400 |
| TTTGTTCCTA CTGTGTTAAT GATAGCTTGA ATCTCTCCTT TCGGTACAAA ACTTGGGATA | 2460 |
| TCCTTATCGT CATCACCAAT CTTACCAACA GGTGGGCTT GCAAGCCATG ACTAGCAAAG | 2520 |
| AAGGCTGCTG TCCCTTCTGT CGCAAGGATT CCATAACCAA TGTTTTGGAA ACGACGAGCC | 2580 |
| AAGTTCAAGG CTCTTCTTT GGCATCATCA GCGATGGTAA AGACGACATT ACCAAAAGTT | 2640 |
| GGCAAGTGTA GATAAGAAGC TTCAAAGGCT TTATAGAGAG CTTTTTCCAA AGTAGCATCA | 2700 |
| GAACCCATAA CTTCACCTGT TGACTTCATT TCAGGACCGA GCAAGCTGTC TACCTTAGCT | 2760 |
| AGTTTGGTAA AGGAGAAGAC AGGTGCCTTG ATATGAACAC GGGTGCTTTC AGGGTAAAGT | 2820 |
| CCATTTTGGT AGCCAAGTTC TGATAAACTT TGACCAAGAA TGAGTTTGGT CGCTACTTGA | 2880 |
| GCCATAGGAA TATTGGTTAC CTTAGATAGG AATGGAACAG TACGGCTGGC ACGTGGATTG | 2940 |
| ACCTCAATAA CGTAGACTTT TTCATCCTTG ATAACAACT GGATGTTTAT CATTTCAAGG | 3000 |
| CAGTGAAGAC CGATTGCTAA GCGTTTGGTG TAGTCTGCGA TGGTCTCCTG AACCTTTTGC | 3060 |
| GACAAGGTTT GTGGTGGGTA AACAGCCATT GAGTCACCTG AGTGGACACC AGCACGTTCTG | 3120 |
| ATATGCTCCA TGATACCAGG AATGAGTACA TTTTACCAT CTGAAATGGC ATCAACTTCG | 3180 |
| CACTCTTGCC CAACGATATA AGAGTCGACA AGAACTGGGT GGTCTGGACT AGCCTTAACA | 3240 |
| GCAGTTTCGA TGTAAGAAGC AAGGTCTTCT TCGTTTTCAA CGATTTCCAT GGCACGTCCA | 3300 |
| CCAAGTACAT AAGATGGGCG GACAAGAAGT GGAAGCCAA TCTTGCGAGC TGCAAGAGCT | 3360 |
| GCTTCTTCTT CATTGGTAGC CGTTTGTCTT GGTGGCTGTG GAATATCCAA TTCTTTGAGA | 3420 |
| GCTTGCTCGA AGAGGTCAGG GTCTTCGGCA CGATCTAGGT CAGCAACCTG TGTACCAAGG | 3480 |
| ATGGTCACAC CTGCTTTTGC CAATGGCTCC GCAAGGTTGA TGGCTGTTTG ACCACCGAAC | 3540 |
| TGAACGATAA CTCCCTTTGG TTGTCCAAG TCAATGACGT TCATAACATC TTCGAATGTC | 3600 |
| AATGGCTCAA AGTAAAGCTT ATCTGATACA GAGAAGTCTG TTGAAACGGT CTCTGGGTTT | 3660 |
| GAGTTCATGA TGATAGCTTC ATAACCAGCT GCCTGGATAG CCTTAACAGA GTGAACGGTT | 3720 |
| GCGTAGTCAA ACTCAACCCC TTGACCGATA CGGATTGGAC CTGAACCTAG GACAAGTACA | 3780 |
| GATTCCTTAT CAGATCTGAT AGATTCATTT TCCCAACCAT AGGTGAATA GAAATATGGC | 3840 |

784

| | |
|--|------|
| GTTTCGGAGT CGAACTCTGC CGCACAAAGTG TCTACCATCT TATAAACTGG AACAACTCTG | 3900 |
| TTTTCCAAGC GAAGTTGGCG AACTTTATCA TCAGTCGTTC CCCAGAGTTC AGCAATCTTA | 3960 |
| CGGTCTGAAA AACCATTAAAG TTTGGCTGTT TTCAAAACTT CTAAATCTTG TGGATGAGCA | 4020 |
| CCCAATTCTT GCTCAATTTC AAAGATATGC AAGAGTTTAT CAAGATAGAA GATATCAATT | 4080 |
| TTTGTAAGCT CTGCAATTTC TTCAGGTGTG TAGCCACGAC GAATGGCTTC TGATACGTAG | 4140 |
| AAGAGACGGT CATCTTGGGC TTTGACAACC TTTTCAATCA AGGCATCATC AGAAACTGCT | 4200 |
| GCAAGTTCAG GTATTTTCATT GTGGTGCACC CCAATTTCAA GGGAGCGGCA GGCCTTGAGA | 4260 |
| AGAGATTCCT CGATGTTACG ACCGATTGCC ATGACTTCTC CAGTCGCCTT CATTTGTGTA | 4320 |
| CCGAGACGGC GTTCACCCCTT TTCAAACTTG TCAAATGGGA AACGTGGAAT CTTAGCAACT | 4380 |
| ACGTAGTCAA GGGCTGGTTC AAACATGGCA TAGGTTGAAC CTGTAACGGT GTTTATAACC | 4440 |
| TCATCCAAGG TCAAACCTAC TGCAATCTTG GCAGCCAAC TAGCAATCGG ATATCCTGTC | 4500 |
| GCTTTAGAAG CAAGGGCTGA CGAACGTGAT ACACGAGGGT TTACTTCGAT AACATAATAC | 4560 |
| TTGAAGCTGT TAGGATCAAG AGCTAGCTGA ACATTACATC CACCTTCAAT CTTGAGGGCA | 4620 |
| CGAATAATGC TCAAGCTCGC ATCACGAAGC ATTTGGTTTT CATAGTCTGA CATGGTTTGC | 4680 |
| GCAGGGGCAA ATACAATGGA ATCCCCTGTG TGAATCCCAA CTGGGTCAA GTTTTCCATG | 4740 |
| TTACAAACAA CCAAGGCATT GTCAGCTGAG TCACGCATCA CTTCGTATTC AATTTCTCTG | 4800 |
| AAACCGGCAA TCGAACGCTC AATCAAACAT TGGGTAACAG GTGACAATT CAAACCATT | 4860 |
| TCAGTGATTT CACGCAATTC TTTCTCGTTG GCACACATAC CACCACCAGT ACCACCAAGG | 4920 |
| GTAAAGGCTG GACGAACGAT GACTGGGTAG CCAATTGTCG CTGCAAAGGC AACTGCTTCT | 4980 |
| TCTACTGTGT TAACAATTTT AGATTCTGGA ATGGGTGTT CAAGCTCTTC CATCAATTGT | 5040 |
| TTAAAGAGGT CACGGTCCTC CGCTTGGTCA ATGGCAGATA ATTTGGTACC CAGAAGTTCA | 5100 |
| ACGCCAAGCT CGTCTAGGAT ACCATTTTGA GATAATTCCA TGGCCATGTT GAGACCTGTC | 5160 |
| TGACCACCGA GTGTTGGTAG CAAGGCATCT GGACCTTCCT TACGAAGAAT ACGTGTCACA | 5220 |
| AACTCAAGTG TAATCGGTTT AATGTAAACC TTGTCAGCAA TTTCTTGTG CGTCATGATG | 5280 |
| GTGTCAGGAT TTGAGTTAAC CAAAACAACC TCATAACCTT CCTCTTTCAA CGACAAGCAA | 5340 |
| GCCTGAGTCC CAGCGTAGTC AAACCTAGCA GCCTGACCAA TAATAATCGG ACCAGAACCA | 5400 |
| ATCACCATAA TTTTGTGAAT ATCAGTACGT TTAGGCATAT ATAAGATATT AAGGGTGTCA | 5460 |
| AGCGGACAAA GCTAAATAG GAGTTATGAC GAAGAACTGT CAGTTCTAGG AATAACTATC | 5520 |
| TTTTTAGCAC CGTCCGTAGC CCGTATTCAG TTCAGCAAAT ACGGAGCACC CTTCTCCTTT | 5580 |
| CTATTCGTG CCTCTCAGGG CGACATTAAA TAAGATACAA AGGACGAATA GAAAGCGATT | 5640 |

785

| | |
|---|------|
| GAATTTTAGG AAATCAAGGA AGGATTGACA ATCCAAGTTG GTTCTCTAC ATTCTGAGCT | 5700 |
| TTCCGTCCGT GTTCAGTTAC ATAAATTCTC CGACGAGCTT TTAGTCTGTC TTAGTTTGAT | 5760 |
| TGTTTAAAAA CTTCATCAT CTCGATAAAC TCGTCAAATA GGTAGCTAGC GTCGTGTGGC | 5820 |
| CCAGGAGCTG CATCTGGGTG GTATTGAACA GAGAAAGCAG GTTGGTATCT GTGGCGCACA | 5880 |
| CCTTCCACTG ACTTGTCATT GATTCTCTCG TGGGTAATAA TCAAGTGCTC TGGCAAATCC | 5940 |
| TCGCGGCTGA CTGCATAACC ATGGTTCTGG CTGGTGAAGT CTACTCGTCC TGTTCGATT | 6000 |
| TCACGTACCG CATGGTTGAA TCCACGGTGG CCAAACTTCA TCTTATAGGT CTTAGCCCCG | 6060 |
| TTTGCCATTG CAAAGAGTTG GTGTCCCAT ACAAATACCA AGATTGGAAT TTTTCCTTGT | 6120 |
| ACACCGCGAA TCATGTCGAG TGCTTGTTGA ACGTCTCTG GGTACCTGG ACCATTGAC | 6180 |
| AACATAACTC CGTCAGGATT GAGATGGAGA ATTTCTTCAG CCGTTGTCGA ATAAGGAACA | 6240 |
| ACTGTCACGT TACAGTTGCG TTTAGAAAGT TCACGTAGGA TTGAGTGCTT GAGACCAAAG | 6300 |
| TCCACTAGCA CCACGCTCAA ACCAACTCCT GGAGCTGGAT AAGACGTTT AGTAGAAACC | 6360 |
| TGTTTGATAT TGTCTGTCGG TAAAACTGTT GCTTGGAGCT GGTCCGTCAC ATGGTCCATA | 6420 |
| CTGTCCCCAA CATGGGTCAA GGTGCGACGC ATAGTACCAT GCTTACGGAT AATCTTGSTA | 6480 |
| AGAGCACGCG TATCAATTCC TGAAATCCCT GGAATTTTCT TGGCTTTCAA AAATTCATCC | 6540 |
| AAGGTCATTT GGTGCGCCA GTTGCTAGCT CTACGCGCTT CTTCAAAAAC AACGACTCCC | 6600 |
| TTACAAGTTG GAATAATGGA TTCATAATCA TCACGATTAA TACCATAATT TCCTACCAA | 6660 |
| GGATAAGTAA AGGTCAAGAT TTGTCCATTA TAAGACTGGT CTGTAATGGA TTCTTGSTAG | 6720 |
| CCGGTCATCC CTGTATTAAA GACGATTTG CCTGTTACAT CAATATCTGC TCCGAAGGCC | 6780 |
| TTGCCTTCAA AAAGTGTGCC ATCTTCTAAT ACTAGAATTC TTTTGTGCAT ATTTTCACCT | 6840 |
| CTCGTGGACG CTCAGTGGCG TCTTTTAACG TCTTGTTTT TAGTTGGCGT TTCTACTCGC | 6900 |
| TAGTACGGAT TCTAAGATTG CCATTCGAAC AAAGACACCA TTGGTCATTT GTTGACAAT | 6960 |
| CCGTGATTTT GGTGCTTCAA CCAAGTGGTC TGCTATTTCT ACATCACGAT TGATTGGAGC | 7020 |
| TGGGTGCATG AGGATTGCTG TTTCTTTCAA ACGATCGTAA CGTTCTTGAG TCAAGCCATG | 7080 |
| TTGGGCATGG TAGTCTTCTT TTGAAAATAC AGCTCCACTA TCATGGCGTT CGTGTGCAC | 7140 |
| ACGGAGAAAC ATCATGACAT CAACCTGATC AATGATTTC TCAATGGTTA CAACTGTCC | 7200 |
| ATAGTCTGCA AACTCTTGAC TTCTCCATTC CTCAGGTCCA GCGAAAAAGA GTTCAGCTCC | 7260 |
| CAAGCGTTTC AAAATCTGCA TATTGGATTT GGCAACGCGT GAGTGGTCCA AGTCACCTGC | 7320 |
| AATAGCAACT TTAAGACCCT CAAAGTGGCC AAATTCCTCA TAAATGGTCA TCAAATCAAG | 7380 |

| | | | | | | |
|------------|------------|------------|------------|------------|-------------|------|
| 786 | | | | | | |
| CAAGCTCTGG | CTAGGGTGTT | GGCCCCAACC | ATCTCCACCA | TTGATGATGG | AAGTCGTAAT | 7440 |
| CGTTGGACTA | GCAATCAATT | CTCTATAGTA | GTCGACCTCT | GGATGGCGAA | TCACACAGAC | 7500 |
| ATCCACTCCT | AAAGCAGACA | GAGTCAAAAT | GGTGTCTATA | AGTGTCTCAC | CCTTATTAAAC | 7560 |
| CGAGCTAGTC | TTCCATCAA | AGTCAAGTCG | TTCCAATCCA | AGTTTAATCT | CTGCGACTTC | 7620 |
| AAAGGACTTA | TGTGTCCGTG | TAGAATCCTC | AAAGAAGAGA | TTGGAAACAA | TCGGATGGTC | 7680 |
| TTCATAGGGA | AGCTGGGCTC | CATTTTTTAA | CTCAATTCCT | CGCTTGATCA | ATTTCATTAC | 7740 |
| TTGATCGACA | GTGAGGTCTT | CCATGGACAC | CACATGGTTC | AATGCTTGTT | GATTTTCTGA | 7800 |
| CATGGCTACT | CCTTTAACTT | TCTAAGCTTC | TTCAGTAATC | AGAACTCTGT | CTTGGTCATC | 7860 |
| AAGTTCTGTC | ATCTCTACGA | TGATTTCTTC | AGAACGACTG | GTTGGGATAT | TTTTTCCAAC | 7920 |
| GTAATCTGGA | CGGATTGGCA | ATTCTCTATG | TCCACGATCG | ACTAGAACTG | CTAAACTCAC | 7980 |
| ACGCGCAGGA | CGACCATGAC | CGACAATATT | ATCAATAGCA | GCACGGATGG | TACGACCTGT | 8040 |
| ATAGAGCACA | TCATCCACCA | AGATAACTTC | GCGGTCTGTC | ACATCGACAG | AAACCAAAGA | 8100 |
| AGTATCTTCT | CCACTTTTAA | CATCATCAGC | GAAAGGTTTA | GTATCCAATT | CCACAACAGG | 8160 |
| AACTGAAAGA | TTTTCTAACT | GCTTCAAACG | TTCTTGGATT | CGGTGGGCAA | TAAAGACACC | 8220 |
| ACGAGTTTAA | ATACCAGCCA | AGACGATCTT | ATTCAAATCT | TTGTTGCGTT | CGATAATCTC | 8280 |
| ATAAGTAATA | CGCGTAATCG | CTCGTTTGAC | GGTCAATTCG | TCTACAACTT | CTTTTGTTTT | 8340 |
| CATGACAAAC | CTCCAAAAAG | AAAAGTCTCC | TTAAACAAGG | AGACTTGAAA | TTTATAGCCA | 8400 |
| AGCGAGCCCT | ACTGCACACA | GTATAGACTT | CACCCTTCTA | CTTTATCGCG | CTCCTTGCCT | 8460 |
| GCCTCACGGG | ACAGGTTTAA | AGGAATATTT | AGTTATCATT | TACTATAGCA | CAAAGCATGC | 8520 |
| TTAAAATCAA | GCAAAAAGTT | TCAATGTAGC | ATCTTACAAA | TTGCTAAAAT | CATATAATTG | 8580 |
| TGGGTACTGG | TCACACTCTG | GATTTTTTGG | ATGGCAAATG | GCTCTTCCAA | AATAAATCAT | 8640 |
| GGCCTGATGG | GCAGCTAACC | ACTGCTCAGG | CGGCAAGATA | TCCATGACCC | GCTTTTCCAC | 8700 |
| CTCAAGTGGC | GTGCTGATT | TTTTGACAAT | ATCGTGGTGT | TTGCAAATAC | GCTCCACAAG | 8760 |
| AGTATCCACT | GCAAAGGCTG | GAATTCCAAA | TCCTACACTC | ATGACAACAT | TGGCTGTCTT | 8820 |
| GCGACCAACA | CCTGCCAAAC | TCTCCAATTC | TTCACTGTGC | TGAGGGACTT | GACCATCAAA | 8880 |
| ATCGTCTAGT | AACTGTTGGG | CACATTTTTT | AAGGAATTTA | GCTTTATTCC | GATACAATCC | 8940 |
| CAAGCGAGAA | ATATGTGAAG | CAATCTCACT | CTCTGTCGCT | ACAGACATAG | CTTGGGGTGT | 9000 |
| TGGAAAGGCA | ACAAAGAGAC | CTGGTGTGGC | CTTATTTACC | GCTGCATCTG | TCGTCTGGGC | 9060 |
| TGATAACATG | ACCGCAACCA | GGAGTTCAAA | ATGATTGGTA | AAATCAAGAC | TAGGCTTGGC | 9120 |
| ATCTGGGAAG | AGGGCAATGA | TTTCTTCTAG | CACCTTTCGT | GCTCGTTTTT | TTGACAAGAC | 9180 |

787

| | |
|---|-------|
| CATTATTCAT CTCCGTCAAA TAGTCCTTGT AAGCCAGCAA AAGGACTGTT TTCTTCTTTC | 9240 |
| TTTACTGCTT TTTGAGCTTG GTATTCTTCC TCTGTCATGA TTTGCCAGTC ATTTCTCGAG | 9300 |
| ATAAATCCTT GACCAGCTTC TTCTTCAGCC GTCAAGACCT TGATAGGAAT GTTTAGCAGG | 9360 |
| ATATTGTCTG ATACACTCTC AGCAAGGTCA AGCTCCCCAT TTTCGATGGG CAAGACCAAG | 9420 |
| TCATCATCTA AAACCTTCTG ATCTAGCTGG TTAGTTGCGC CTTCCATGAA AACTTCCGTG | 9480 |
| ACTGGATAAG ATTCAACTAA CTCAACTGGC TCCATACTGC GACTCGACGC AAGAACAATG | 9540 |
| GTATAAGATA GTTGATAATC TAAGAAATAC ATACGGTCTT CATATTGTAC TTCCCAACT | 9600 |
| GCAAGGATAT CTTTACATC TAAAATTTCT TGATTACGTG CACGCAGGTC ATCAACTAAA | 9660 |
| TCTAACGTTT GTTCAAAGTT CAAACCTTCA GACTGCTTAC GAATTTCTTG AATATTTAAT | 9720 |
| TTCATACTTC CTCCATAAAG ATTTACTCTC TTGATTATAC CATGAAAAGG CTACAAATCA | 9780 |
| GCACACCAAA CTTTGTAATT AAAATTCAAA ATTTTAACAT ATTTACTATG ATAGTTTAT | 9840 |
| TTTTTAGTGC TATACTATAG GGAAAGAGTA CATCAGATCA AGGAGGATGC TCACATGGAA | 9900 |
| GACAAGAAAC TCATTCAACT CCTATCCAAG TTAAATAAAA GCTACCAAAA CTGTAAACAG | 9960 |
| GGTACGGCAG ATGATATTCG ACTACAAGAG CTGCTAAACA CTACTATGCA AGAGCTCAAA | 10020 |
| AAAACGGAAC AGTTGAACAA CAGTATCTTA ATTGATCTTG AGAAATTTTA CCAACCTACC | 10080 |
| AGTCTTCTGA TTGGACTGGG TAGCCTAAAA CTAAACGATC AAGCACGCAC TGCTTGGCGA | 10140 |
| AACTATGATA AATTCCATTA CGATCATGTC AAACACGTAC TAAGTCTCTA TGGACCTGTT | 10200 |
| TTTGAATTTT AGAGCATAGA ATTTCCAGTT TTCTGTTGAC AAAATTTCTT TAAAGGTATA | 10260 |
| ATATAAGAT ACTAATACTC GGAGGTAAGG GAGACATGAA CAACTAAGTC TATCAAATAA | 10320 |
| AGAACCTTTA TTTAGTAGAT CTTGTTTTTG TCTCTTTTTG TGTGCTCTTT TATGCTCTTT | 10380 |
| TTCTGGCATG TTAATAGAGT TTTTTTGACA TAGACTTTGG GCTCTACTAG GTAAAGTAGA | 10440 |
| GCTTTTTGTT ATGCACTATG AACATTCTAG AAAGGGAAAT CATATGATAA AAATCAATCA | 10500 |
| TCTAACCATC ACACAAAACA AAGATTTACG AGATCTTGTA TCTGACCTAA CCATGACCAT | 10560 |
| CCAAGACGGG GAAAAGGTTG CTATTATTGG TGAAGAAGGA AATGGCAAAT CAACCTTACT | 10620 |
| TAAAATTTTA ATGGGGGAAG CTTTGTCTGA TTCTACTATC AAGGGAAACA TCCAATCTGA | 10680 |
| CTATCAGTCA CTGGCCTACA TTCCTCAAAA AGTCCCTGAG GACCTAAAAA AGAAAACCTT | 10740 |
| ACACGACTAC TTCTTTTATG ATTCATTGA TTTAGACTAC AGTATCCTCT ATCGTTTGGC | 10800 |
| GGAGGAATTG CATTTTGATA GCAATCGTTT CGCAAGTGAC CAAGAGATTG GCAATCTATC | 10860 |
| AGGGGGCGAA GCTTTGAAAA TTCAGCTTAT CCATGAGTTA GCCAAACCCT TTGAGATTCT | 10920 |

788

| | |
|--|-------|
| ATTTTGTAGAT GAACCTTCAA ATGACCTAGA CCTTGAGACA GTTGATTGGC TAAAAGGCCA | 10980 |
| GATTCAAAAG ACCAGGCAAA CCGTTATTTT CATTTCCCAT GATGAAGACT TTCTTTCTGA | 11040 |
| AACGGCAGAC ACTATTGTTT ACTTGCGACT GGTCAAACAC CGTAAAGAAG CGGAAACGCT | 11100 |
| AGTAGAGCAT TTAGACTATG ATAGCTATAG TGAGCAGAGA AAGGCTAATT TTGCCAAACA | 11160 |
| AAGTCAGCAA GCTGCTAACA ACCAAAGAGC CTACGATAAA ACCATGGAAA AACATCGGAG | 11220 |
| AGTTAAGCAA AATGTAGAAA CTGCGCTTCG AGCTACCAA GATAGTACTG CCGGTCGCCT | 11280 |
| ATTGGCTAAA AAGATGAAAA CTGTCCTCTC ACAAGAAAA CGCTACGAAA AGGCAGCTCA | 11340 |
| GTCCATGACT CAAAAGCCAC TTGAAGAGGA ACAAATCCAA CTTTCTTTT CAGACATCCA | 11400 |
| ACCATTACCA GCTTCTAAAG TCTTAGTCCA ACTGGAAAA GAAAATTTGT CCATTGACGA | 11460 |
| CCGAGTTTGT GTTCAAAAAC TACAACCTAAC TGTCCGTGGC CAAGAAAAA TCGGTATTAT | 11520 |
| CGGGCCAAAT GGTGTGGGA AATCAACTCT GTTAGCCAAG TTACAGAGAC TTCTGAATGA | 11580 |
| TAAAAGAGAG ATTTCACTTG GTTTATGCC ACAAGATTAC CACAAAAAAC TGCAATTGGA | 11640 |
| TTTATCCCA ATAGCCTATC TCAGTAAAC TGGGAAAAA GAGGAACTAC AGAAAATCCA | 11700 |
| ATCTCACCTA GCTAGTCTCA ATTTAGTTA TCCAGAAATG CAGCATCAA TTCGCTCCTT | 11760 |
| ATCTGGCGGA CAACAGGGAA AACTCCTGCT TTTGGATTGA GTCCTGCGCA AACCAACTT | 11820 |
| TCTCCTGCTG GATGAACCA CACGAACTT TTCTCCCTCT TCTCAACCCC AAATCAGAAA | 11880 |
| ACTCTTTGCT ACCTATCCAG GCGGTCTCAT CACTGTTTCG CATGACCGTC GTTCTTAAA | 11940 |
| AGAAGTCTGC TCGATCATCT ATCGCATGAC AGAACACGGT TTGAAGCTAG TTAATTTAGA | 12000 |
| AGATTATAA ATTTGCAACA TAGCAAAAT CCAGAGACGA CCTCTGGATT CTTTACATC | 12060 |
| TCTTTTAAAC GTTCAATCCG TTCTGAGATA GGTGGGTGGG TATAAAGAG TTTTGGAAC | 12120 |
| CCCCACCTT TCTTAGGATC ATTGATATAA AGGGCACTGC TAGCATCATC GACGTGGCGA | 12180 |
| CTCATAGGTT TGCTATTGTC CAACTTATCT AGGGCATTAA TCATTCCCTG GGGATTGCGA | 12240 |
| GTCAGCTCGA CACTAGATGC ATCTGCCAGA AATTCCTCT GACGAGAAAT AGCGAGCTGA | 12300 |
| ACCAAGGTTG CAGCGAGAGG TGCCAGTACA ATAGCTAGTA GGGAAACCAC TAGCATAATG | 12360 |
| ATTTCAAGAC CATTTCCATC TCGGTCATCA TCACTTCGTC TGCGACCTGC TCCACCCAC | 12420 |
| CACATCATAC GACCTGCCAT ACTAGAAAGC ATGGTGATAG CACTAGCAAG GGCAACTGCA | 12480 |
| ATAGTCGAAA TACGGATATC ATAATTACGA ATATGACTGA CTTCATGTCC CATAACAGCT | 12540 |
| TCTAGTTCTT CACGATTCAT GATAGCTAGT AGACCTGAAG TCGCAGCAAC AGCCGCATTT | 12600 |
| TGAGGATTAG AACCTGTCG AAAGGCATTT AAGGCTGGAT CATCAATGAT GAAAACACGG | 12660 |
| GGCATAGGAA TCTGAGCGAC CAGAGCCATA TCTTCCACTA CATGGTAGAG GTCTGGTGCC | 12720 |

789

| | |
|--|-------|
| GTTTGCTCAT CCACCTCAGC CGCTCCATTC ATGGACATGA CAATCTCTGT CGATTGAAAA | 12780 |
| ATCATAGACA AAGCGTAGAT AAAGCCGATA ATCAGTGCAA TAACCAAACC ACCAAGTCCA | 12840 |
| GATCTTATAA AGAGATAACC AACCGCATAA CCAACAAGAG CTAAGAGTAG GAAAAATACC | 12900 |
| AGCAACAAAA TCCAGGTTTT TCGTTTATTG CTTGCAATTT GATCAAACAA CATCTTAGTC | 12960 |
| ACCTAAACCG CTAAATCAA CTTTAGGAAC CGACTTTTCC TCTTCAGGTG TTTGAAGGAA | 13020 |
| ATCTGCCGCT TTAAATCCAA ACATTCCAGC GATAATATTG CTCGGGAAAG TTTCTAATTT | 13080 |
| TACATTGTAG TTGCTGACAA CACTGTTATA GAGTTGACGA GAGTAAGAAA TTTTATTTTC | 13140 |
| TGTGTTGTGTC AACTCCTCTT GCAATTTAAC AAAGTTAGCA CTAGCTTTCA AATCTGGATA | 13200 |
| GCTTTCTGCA ACTGCAAAAA TACCTGAAAC CTGACGAGTG AGGGCATCAC TGGCTTTCAT | 13260 |
| AGCTTCTGCT GGTGAAGTCG CTGCCGCCAC TTGGTTACGT AGTTCTGCCA CCTTTTCAAG | 13320 |
| GGTAGAACCT TCATATTTGG CATAACCTTT TACAGTCTCA ATCAAGTTTG GCAAGAGGTC | 13380 |
| ATTGCGACGT TTCAACTGAA CATCAATCTG ACTCCAAGCC TCCTTGGTTT GCATACGATT | 13440 |
| TTTAACCAAA CCGTTATAGC TAACAATCAC AAAAATAACA ATAAGAGCGA TAACTCCAAG | 13500 |
| AATAATCCAA GTCATAATAT AAGTCCTTTC TGCTTTTAGA TTAGTACCAG TATATCAAAT | 13560 |
| TTTCTATGAT TGTGGTAAAA TAAGATGATA CTAAAGAAGG AAATAACTAT GAAACCAAAA | 13620 |
| ACATTTTACA ACTTGCTTGC CGAGCAGAAT CTTCACCTTT CGGACCAGCA AAAAGAACAA | 13680 |
| TTTGAACGTT ATTTTGAGCT CTGGTTCGAG TGGAATGAGA AGATTAATTT GACGGCGATT | 13740 |
| ACGGACAAGG AAGAAGTTTA TCTCAAACAT TTTTACGATT CGATTGCACC CATTTCTCAA | 13800 |
| GGTTTGATTC CCAATGAAAC TATCAAACCT CTTGATATCG GGGCTGGGGC AGGATTTCTT | 13860 |
| AGTCTACCAA TGAAAATTCT CTATCCGGAG TTAGATGTGA CCATTATTGA TTTCACTCAAT | 13920 |
| AAGCGCATCA ACTTCCTACA ACTCTTGGCT CAAGAACTGG ATTTGAACGG AGTTCAATTT | 13980 |
| TACCACGGAC GTGCCGAAGA TTTTGCCCAA GACAAGAACT TCCGTGCTCA ATATGATTTT | 14040 |
| GTAACAGCTC GTGCGGTTGC CCGTATGCAG GTCCTATCTG AATTGACTAT TCCCTACCTT | 14100 |
| AAGGTTGGTG GCAAACCTATT AGCACTCAAG GCTAGCAATG CGCCTGAGGA ATTATTAGAA | 14160 |
| GCTAAGAATG CCTCAATCT CCTTTTATG AAGGTCGAAG ACAATCTCAG TACGCCCTAC | 14220 |
| CGAATAGAGA TCCGCGCTAT ATCACAGTGG TAGAAAAGAA AAAAGAAACA CCAAATAAAT | 14280 |
| ATCCACGTAA GGCTGGTATG CCAAATAAAC GCCCACTTTA AATTMTTTAG TAAACAAATG | 14340 |
| TTTACAAAAT CAGCCTCGCT CTTTATTATTC TAGGCTCGGG AAAAAATGAT TTACAAAATC | 14400 |
| AGCCTCGCTC TTTTATTCTT AGGCTCGGGA AAAAATGATT TACAAAATCA TTTTMTCTG | 14460 |

790

| | |
|---|-------|
| CTATACTATC CTAAGCAAAG GTTTTTAATG TCATCCCGTG AGGTGACGAA GACGCAGAAA | 14520 |
| TATTTAAAAC TCTTTAAAT CTAAATTTTA AAGAAGTCTT ACTCTGAGGG CCTATTGCTG | 14580 |
| TAAAATAATG GGCTCTTTTT TGATGCCCAA AAGTGAGGTT TATATGAAAC AAGAATCAAC | 14640 |
| TGTTGATTG TTAC | 14654 |

(2) INFORMATION FOR SEQ ID NO: 107:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 6405 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 107:

| | |
|--|------|
| AGAAAAATCT GCTTTACAGA AAATAAAAT AATAGGAGAA AATCTATGTC AGATTTGAAA | 60 |
| AAATACGAAG GTGTCATTCC AGCCTTCTAC GCATGTTATG ATGATCAAGG AGAAGTAAGC | 120 |
| CCAGAACGTA CGCGTGCCTT GGTCAATAC TTCATTGATA AAGGTGTTCA AGTCTTTAT | 180 |
| GTCAATGGTT CTTCTGGTGA ATGTATCTAC CAAAGCGTTG AAGATCGCAA GTTGATTTTG | 240 |
| GAAGAAGTCA TGGCGGTAGC AAAGGTAAAT TGACCATTAT TGCCCATGTT GCTTGCAATA | 300 |
| ATACTAAAGA TAGTATGGAA CTTGCTCGCC ATGCTGAAAG CTTGGGAGTA GATGCTATTG | 360 |
| CAACGATTCC ACCAATTAT TTCCGCTTGC CAGAATACTC AGTTGCCAAA TACTGGAACG | 420 |
| ATATCAGTTC TGCAGCTCCA AACACAGACT ACGTGATTTA CAACATTCCT CAATTGGCAG | 480 |
| GGGTTGCTTT GACTCCAAGC CTTTACACAG AAATGTTGAA AAATCCTCGT GTTATCGGTG | 540 |
| TGAAGAACTC TTCTATGCCA GTTCAAGATA TCCAAACCTT TGTGAGCCTT GGTGGAGAAG | 600 |
| ACCATATCGT CTTTAATGGT CCTGATGAGC AGTTCCCTAGG AGGACGCCTC ATGGGGGCTA | 660 |
| GGGCTGGTAT CGGTGGTACT TATGGTGCTA TGCCAGAACT CTTCTTGAAA CTCAATCAGT | 720 |
| TGATTGCGGA TAAGGACCTA GAAACAGCGC GTGAATTGCA GTATGCTATC AACGCAATCA | 780 |
| TTGGTAAACT CACTTCTGCT CATGGAAATA TGTACGGTGT CATCAAAGAA GTCTTGAAAA | 840 |
| TCAATGAAGG CTTGAATATT GGATCTGTTC GTTCACCAT TACACCAAGT ACTGAAGAAG | 900 |
| ATCGTCCAGT TGTAAGAAGC GCTGCTGCCT TGATTCGTGA AACCAAGGAG CGTTCTCTCT | 960 |
| AATCTAAAAG GAGGTATTTA TGACATATTA CGTTGCAATT GATATCGGTG GAACCAACAT | 1020 |
| CAAGTATGGT TTGGTTGATC AAGAGGGGCA ACTTCTTGAA TCGCATGAAA TGCCAACTGA | 1080 |
| GGCGCATAAG GGTGGACCTC ATATCTTACA AAAGACCAA GATATCGTAG CTAGTTATTT | 1140 |
| AGAAAAAGGC CCAGTAGCAG GTGTTGCCAT ATCTTCTGCT GGGATGGTGG ATCCGGATAA | 1200 |

791

| | |
|---|------|
| GGGTGAGATT TTCTATGCTG GGCCGCAAAT CCCTAACTAC GCAGGCACCC AGTTCAAAAA | 1260 |
| GGAAATCGAA GAAAGCTTTA CTATTCCTTG TGAGATTGAA AATGATGTCA ACTGTGCAGG | 1320 |
| TCTTGCTGAG GCAGTATCTG GTTCAGGCAA GGGAGCAAGT GTGACACTTT GCTTGACCAT | 1380 |
| TGGAACCGGT ATCGGTGGTT GCTTGATTAT GGATAGGAAA GTCTTCCATG GTTTTAGCAA | 1440 |
| TTCAGCCTGT GAAGTCGGT ATATGCATAT GCAGGATGGA GCTTTTCAAG ACTTGGCTTC | 1500 |
| TACAACAGCT TTAGTGAAAT ATGTAGCTGA AGCCCATGGA GAAGATGTTG ATCAGTGGA | 1560 |
| TGGCCGTAGA ATTTTCAAAG AAGCCACTGA AGGAAACAAA ATCTGCATGG AAGGTATTGA | 1620 |
| CCGTATGGTT GACTATCTAG GAAAAGGTCT GGCAAATATT TGCTACGTTG CCAATCCAGA | 1680 |
| AGTGGTTATT CTTGGTGGTG GTATCATGGG GCAAGAGGCT ATCCTCAAAC CTAAGATCCG | 1740 |
| TACAGCCTTG AAAGAGGCTT TGGTACCAAG TTTAGCAGAA AAAACACGAT TAGAATTTGC | 1800 |
| CCATCACCAA AATACAGCAG GGATGTTGGG TGCATATTAT CATTTTAAGA CAAAACAATC | 1860 |
| CTAGTTTGGC TCAGCCAAAC TAGGATTTTC TTACACGTTT TTGTCTACGA TAGCCGTTGA | 1920 |
| GTTCCTTATT TTCCAGTAG CTATTAAAGA TTTTTCCTT GCTTTCGCGA TTGATTTCCA | 1980 |
| AAAAGTAGGC ATAAATCAA TCGATAAGA AGAGCATAGG AAGTTGAGCG GATATTCGTT | 2040 |
| GGATATAGGA GGGTTGGCTG TGGGTGGCTA CAAGAACAGT CTCTGTATAG GTCTGGCTAT | 2100 |
| CTTTATTGGG AACACTTGTA AAGAGTACAG TCTTTCCTT CATCTCCTTA GCATCTAATA | 2160 |
| GACTATCTAA AATAGAAGGA GTTGAGCCTG AAAGTGAGAA GCCCAGTACT AGACAATTTT | 2220 |
| CATCCATGAT GCTGGTTGTC CAGGCAAAGC CGTCTTGGTC TGTCAAAGCT TCGCAGACCA | 2280 |
| CACCTAGTCG CATAAACGT AATTTTATTT CACGGGCGAC GAGGCCAGAA CTCCTGTTC | 2340 |
| CAAAGAAGTA GATACGCTCA GCATCTTCGA TTAGCTGGGC AATTCGTTCT AGTTGGATTT | 2400 |
| CGTCAATCAA GTCTTGTTT TGTTCCTCA TATTGCTATA ACTTCTGAGG ACTCGTTTGG | 2460 |
| TCAGTGGACT GTGCTGGAG ACTTGGTTGG CTTGATTTTC TGCCTGATGT TGGTATTGGA | 2520 |
| AAATAAATTC TCGGTAGCCA GTAAAGCCAC ACTTTTTCAG AAAGCGGGTC AAAGCAGCTT | 2580 |
| GAGAAATATG TAATTTTGG GTGACTTGTT GAGAAGATAA ATCATCTGTA ATCGTTTCAG | 2640 |
| CTTGCAAAAA ATAGCGAGCG ATTTCTTGTT CTAGGTCTGT CATTTCTTCA AAATGTGAAT | 2700 |
| CAATGATAGT TCGATATCT GGTGTGCTCA TAGGGAAAGC TCCTTTACAT GAGTCATACT | 2760 |
| GGAAGACTAG ATCAGAGAAT AGTCACACTT CATTATAACA CATAATATAA GGATAGATAA | 2820 |
| ATAAAAACGC ATCTCTGTTT TAAAAACGAA AAAATCGAAA AAGCTTCTCT CTTTCCATA | 2880 |
| ATTTTCTACT CAAATTGTGG TACAATTAAG AGTAAGATTT TAAGTTAGAA ATGAGACTGA | 2940 |

792

| | |
|--|------|
| TTTGATGAG AAAATTAAAC AGCCATTCGA TTCCGATTCG GCTTAATTTA TTGTTTTCAA | 3000 |
| TCGTCATTTT ACTCTTTATG ACCATTATTG GTCGTTTGTT GTATATGCAG GTTTTGAACA | 3060 |
| AGGATTTTTA CGAAAAAAG CTAGCTTCAG CTAGTCAGAC CAAGATTACA AGCAGTTCAG | 3120 |
| CCCGTGGGGA AATTTATGAT GCTAGTGGAA AACCTTTGGT AGAAAATACG TTAAAGCAGG | 3180 |
| TTGTTTCCTT TACGCGTAGC AATAAAATGA CGGCTACAGA CTTAAAAGAA ACAGCTAAAA | 3240 |
| AGTTACTGAC TTATGTGAGC ATCAGTTCTC CAAATTTGAC AGAACGCCAG CTGGCGGATT | 3300 |
| ACTATTTGGC TGATCCTGAA ATCTATAAAA AAATAGTGGA AGCTCTCCCA AGTGAGAAAC | 3360 |
| GCTTGGATTC AGATGGCAAT CGTCTATCCG AATCAGAACT GTATAACAAT GCGGTCGATA | 3420 |
| GTGTACAAAC GAGTCAACTA AACTATACAG AGGATGAAAA GAAAGAAATC TATCTTTTTA | 3480 |
| GTCAGTTAAA TGCTGTTGGA AACTTTGCGA CAGGAACCAT TGCGACAGAT CCTCTAAATG | 3540 |
| ATTCTCAGGT GGCTGTTATT GCCTCTATTT CAAAGGAGAT GCCTGGCATT AGTATTTCTA | 3600 |
| CTTCTTGGA TAGAAAGGTT TTGGAAACTT CCCTTTCTTC TATAGTTGGG AGTGTATCCA | 3660 |
| GTGAAAAAGC TGGTCTCCCA GCGGAAGAAG CAGAAGCCTA TCTTAAAAAA GGCTATTCTC | 3720 |
| TAAATGACCG TGTAGGAACC TCCTATTTGG AAAAGCAATA TGAAGAGACC TTACAAGGAA | 3780 |
| AACGCTCGGT AAAAGAAATC CATCTGGATA AATATGGCAA TATGGAAAGC GTGGATACAA | 3840 |
| TTGAGGAAGG TAGTAAGGGA AACAATATCA AACTGACCAT TGATTTGGCT TTCCAAGATA | 3900 |
| GCGTGGATGC TTTACTGAAA AGTTATTTCA ATTCTGAGCT AGAAAATGGT GGAGCCAAGT | 3960 |
| ATTCTGAAGG TGTCTATGCA GTCGCCCTTA ACCCAAAAAC AGGTGCGGTT TTGTCTATGT | 4020 |
| CAGGGATTAA ACATGACTTG AAAACGGGAG AGTTGACGCC TGATTCCTTG GGAACGGTAA | 4080 |
| CCAATGTCTT TGTTCCAGGT TCGGTTGTCA AGGCGGCGAC CATCAGCTCA GGTGGGAAA | 4140 |
| ATGGAGTCTT GTCAGGAAAC CAGACCTTGA CAGACCAGTC CATTGTCTTC CAAGGTTTCAG | 4200 |
| CTCCCATCAA TTCTTGGTAT ACTCAGGCTT ACGGTTTCATT CCCTATCACA GCGGTCCAAG | 4260 |
| CTCTGGAGTA TTCATCAAAT ACCTATATGG TCCAAACAGC CTTAGGTCTT ATGGGGCAAA | 4320 |
| CCTATCAACC CAATATGTTT GTCGGCACCA GCAATCTAGA GTCTGCTATG GAGAACTGC | 4380 |
| GTFCAACCTT TGGCGAATAT GGCTTGGGTA CTGCGACAGG AATTGACCTA CCAGATGAAT | 4440 |
| CTACTGGATT TGTTCCCAA GAGTATAGCT TTGCTAATTA CATTACTAAT GCCTTTGGGC | 4500 |
| AGTTTGATAA CTATACGCCG ATGCAGTTGG CTCAGTATGT AGCAACTATT GCAAATAATG | 4560 |
| GTGTTTCGTG GGCTCCTCGT ATTGTTGAAG GCATTTATGG TAATAATGAT AAGGGAGGAC | 4620 |
| TGGGTGACTT GATTCAGCAA CTGCAACCGA CAGAGATGAA TAAGGTCAAT ATATCCGACT | 4680 |
| CCGATATGAG CATCTGACAC CAAGGTTTTT ATCAGGTTGC CCATGGTACT AGTGGATTGA | 4740 |

793

| | |
|--|------|
| CAACTGGACG TGCCTTTTCA AATGGTGCCT TGGTATCCAT TAGCGGAAAA ACAGGTACAG | 4800 |
| CCGAAAGCTA TGTGGCAGAT GGTGAGCAAG CAACCAATAC CAATGCGGTG GCCTATGCCC | 4860 |
| CATCTGATAA TCCCCAAATC GCTGTCGCAG TGGTCTTTCC TCATAATACC AATCTAACAA | 4920 |
| ATGGTGTAGG ACCTTCCATT GCGCGTGACA TTATCAATCT GTATCAAAAA TACCATCCAA | 4980 |
| TGAATTAGAA AGGAAATTAT GCTTTATCCA ACACCTATTG CCAAGTTGAT TGACAGTTAT | 5040 |
| TCTAAGTTAC CAGGTATCGG GATTAAGACG GCTACGCGTC TGGCCTTTTA TACGATTGGG | 5100 |
| ATGTCTGCTG ATGATGTCAA TGAATTTGCA AAAAATCTCC TTTCTGCTAA GAGAGAATTG | 5160 |
| ACATATTGTT CTATTTGTGG ACGTTTGACA GACGACGATC CTTGTTCTAT CTGTACTGAT | 5220 |
| CCGACTCGTG ACCAGACAAC AATTTTAGTT CTTGAGGATA GTAGAGATGT GGCAGCCATG | 5280 |
| GAAAATATCC AAGAATACCA TGGACTCTAT CATGTCCTTC ATGGCCTCAT TTCTCCTATG | 5340 |
| AATGGTATCA GTCCGGACGA TATCAATCTC AAGAGCCTTA TGAATCGTCT TATGGATAGT | 5400 |
| GAGGTTTCAG AAGTGATTGT GCGGACTAAT GCTACAGCGG ATGGTGAAGC GACTTCCATG | 5460 |
| TATCTTTTAC GTTTGCTCAA GCCGGCTGGT ATCAAGGTTA CGCGTCTAGC ACGAGGTCTC | 5520 |
| GCTGTGGGAG CGGACATTGA GTATGCGGAC GAAGTGACAC TCTTACGAGC CATTGAAAAAT | 5580 |
| CGGACAGAGT TGTAAGTGTA GGCAAATTTA CGAACTCCAT TCATTTATAA AAAATCAAAG | 5640 |
| AGGCTGAAAA TCGTTCCTAT CGGCCTCTTT TTGTATAGTG TGATGACTAG GCTCAGGTTT | 5700 |
| AAGTTTAAAA AAACCAAGCA AATATGATAT ACTAAAGAGC GAGTATTCTA GTAGAATTAG | 5760 |
| GACAAATAAT ATGAAACAAA CGATTATCTT TTTATATGGT GGACGGAGTG CGGAACGCCA | 5820 |
| AGTCTCTGTC CTTTCAGCTG AGAGTGTCAT GCGTGCGGTC GATTACGACC GTTTCACAGT | 5880 |
| CAAGACTTTC TTTATCAGTC AGTCAGGTGA CTTTATCAAA ACACAGGAAT TTAGTCATGC | 5940 |
| TCCGGGGCAA GAAGACCGTC TCATGACCAA TGAAACCATT GATTGGGATA AGAAAGTTGC | 6000 |
| ACCAAGTGCT ATCTACGAAG AAGGTGCAGT GGTCTTTCCA GTCCTTCACG GGCCAATGGG | 6060 |
| AGAAGATGGC TCTGTTCAAG GATTCTTGGA AGTTTGTAAA ATGCCCTACG TTGGTTGCAA | 6120 |
| CATTTTGTCA TCAAGTCTTG CCATGGATAA AATCACGACT AAGCGTGTTC TGGAACTGTC | 6180 |
| TGGTATTGCC CAAGTTCCTT ATGTGGCTAT CGTTGAAGGC GATGATGTGA CTGCTAAAAAT | 6240 |
| CGCTGAAGTG GAAGAAAAAT TGGCTTATCC AGTCTTCACT AAGCCGTCAA ACATGGGGTC | 6300 |
| TAGTGTCGGT ATTTCTAAGT CTGAAAACCA AGAAGAATC CGTCAAGCCT TAAACTTGC | 6360 |
| CTCCGATAT GACAGCCGTG TCTTGGTTGA GCAAGGAGTG AATGC | 6405 |

(2) INFORMATION FOR SEQ ID NO: 108:

794

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 11309 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 108:

| | |
|--|------|
| CGAGCTCGGG TACCGGGATT TTAAGGAGTT TGATATGTAT AACCTATTAT TAACCATTTT | 60 |
| ATTAGTATTA TCTGTTGTGA TTGTGATTGC AATTTTCATG CAACCAACCA AAAACCAATC | 120 |
| CAGCAATGTA TTTGATGCCA GTTCAGGTGA TTTGTTTGAA CGCAGTAAAG CTCGCGGTTT | 180 |
| TGAAGCTGTA ATGCAGCGTT TGACAGGGAT TTTAGTCTTT TTCTGGCTAG CCATTGCCTT | 240 |
| AGCATTTGACG GTATTATCAA GTAGATAAGA AAATAATGGG CAGGACTAGG TCTTTGCCTC | 300 |
| TTTTTATTTT TAAAGGATGT TTGAGAAGGT TTTACAGTAA AAGAAAATTA AAAAATCTAG | 360 |
| AAAGAAAATA TGAAAGATAG AATAAAAGAA TATTTACAAG ACAAGGGAAA GGTGACTGTT | 420 |
| AATGATTTGG CTCAGGCTTT GGGAAAAGAC AGTTCCAAGG ATTTTCGTGA GTTGATTAAA | 480 |
| ACCTTGTCCT TAATGGAAAG AAAGCACCAA ATTCGTTTGG AAGAAGATGG TAGTCTGACA | 540 |
| TTAGAAATTA AGAAAAACA TGAGATTACC CTCAAGGGGA TTTTTCATGC CCATAAAAAAT | 600 |
| GGCTTTGGCT TTGTTAGTCT GGAAGGCGAG GAGGACGACC TTTTGTAGG GAAAAATGAT | 660 |
| GTCAACTATG CTATTGATGG TGATACCGTC GAGGTAGTGA TTAAGAAAGT CGCTGACCGC | 720 |
| AATAAGGGAA CAGCAGCAGA AGCCAAAATT ATTGATATCC TAGAACACAG TTTGACAACA | 780 |
| GTTGTCGGGC AAATCGTTCT GGATCAGGAA AAACCTAAGT ATGCTGGCTA TATTCGTTCA | 840 |
| AAAAATCAGA AAATCAGTCA ACCGATTTAT GTTAAGAAAC CAGCCCTAAA ATTAGAAGGA | 900 |
| ACAGAAGTTC TCAAAGTCTT TATCGATAAA TACCCAAGCA AGAAACATGA TTTCTTTGTC | 960 |
| GCGAGTGTTT TCGATGTAGT GGGACACTCA ACGGATGTCG GAATTGATGT TCTTGAGGTC | 1020 |
| TTGGAATCAA TGGACATTGT ATCCGAGTTT CCAGAAGCTG TTGTTAAGGA AGCAGAAAGT | 1080 |
| GTGCTGATG CTCCGTCTCA AAAGGATATG GAAGGTCGTC TGGATCTAAG AGATGAAATT | 1140 |
| ACCTTTACCA TTGACGGTGC GGATGCCAAG GACTTGGACG ATGCAGTGCA TATCAAGGCT | 1200 |
| CTGAAAAATG GCAATCTGGA GTTTGGGGTT CACATCGCAG ATGTTTCTTA TTATGTGACC | 1260 |
| GAGGGGTCTG CCCTTGACAA GGAAGCCCTT AACCGTGCGA CTTCTGTTTA CGTGACAGAC | 1320 |
| CGAGTGGTGC CAATGCTTCC AGAACGACTA TCAAAATGGCA TCTGCTCTCT CAATCCCCAA | 1380 |
| GTTGACCGCC TGACCCAGTC TGCTATTATG GAGATTGATA AACATGGTCG TGTGGTCAAC | 1440 |
| TATACCATTA CACAAACAGT TATCAAGACC AGTTTTCGTA TGACCTATAG CGATGTCAAT | 1500 |

795

| | |
|--|------|
| GATATCCTAG CTGGCGATGA AGAAAAGAGA AAAGAATATC ATAAAATTGT ATCAAGTATC | 1560 |
| GAACCTCATGG CCAAGCTTCA TGAAACTTTA GAAAACATGC GTGTGAAACG TGGAGCTCTC | 1620 |
| AATTTTGATA CCAATGAAGC GAAGATTTTA GTGGATAAAC AAGGTAAGCC TGTTCATATC | 1680 |
| GTTCTTCGGC AGCGTGGTAT TGCCGAGCGG ATGATTGAGT CTTTATGTT GATGGCTAAT | 1740 |
| GAAACAGTTG CCGAACATTT CAGCAAGTTG GATTTGCCTT TTATCTATCG AATTCACGAG | 1800 |
| GAGCCTAAGG CTGAAAAGGT TCAGAAGTTT ATTGATTATG CTTTCAGTTT TGGCTTGCGC | 1860 |
| ATTTATGGAA CTGCCAGTGA GATTAGTCAG GAGGCACTTC AAGACATCAT GCGTGCTGTT | 1920 |
| GAGGGAGAAC CTTATGCAGA TGTATTGTCC ATGATGCTTC TTCGCTCTAT GCAGCAGGCT | 1980 |
| CGTTATTTCGG AGCACAATCA CGGCCACTAT GGACTAGCTG CTGACTATTA TACTCACTTT | 2040 |
| ACCACTCCAA TTCGTCGTTA TCCAGACCTT CTGTTCACC GTATGATTCG GGATTACGGC | 2100 |
| CGTTCTAAGG AAATAGCAGA GCATTTTGAA CAAGTGATTC CAGAGATTGC GACCCAGTCT | 2160 |
| TCCAACCGTG AACGTCGTGC CATAGAAGCT GAGCGTGAAG TCGAAGCCAT GAAAAAGGCT | 2220 |
| GAGTATATGG AAGAATACGT GGGTGAAGAG TATGATGCAG TTGTATCAAG TATTGTCAAA | 2280 |
| TTCGGTCTCT TTGTGCAATT GCCAAACACA GTTGAAGGCT TGATTCACAT CACTAATCTG | 2340 |
| CCTGAATTTT ATCATTTCAA TGAGCGTGAT TTGACTCTTC GTGGAGAAAA ATCAGGTATC | 2400 |
| ACTTTCGGAG TGGGTCAGCA GATCCGTATC CGTGTGAAA GAGCGGATAA AATGACTGGA | 2460 |
| GAGATTGATT TTTCATTCTG ACCTAGTGAG TTTGATGTGA TTGAAAAAGG CTTGAAACAG | 2520 |
| TCTAGTCGTA GTGGCAGAGG GCGTGATTCA AATCGTCGTT CGGATAAGAA GGAAGACAAG | 2580 |
| AGAAAATCAG GACGCTCAAA TGATAAGCGT AAGCATTCAC AAAAAGACAA GAAGAAAAAA | 2640 |
| GGAAAGAAAC CTTTTTACAA GGAAGTAGCT AAGAAAGGAG CCAAGCATGG CAAAGGGCGA | 2700 |
| GGGAAAGGTC GTCGCACAAA ATAAAAAGGC ACGCCACGAC TATACAATCG TAGATACGCT | 2760 |
| AGAGGCAGGG ATGGTCCTGA CTGGAACTGA AATCAAGAGT GTACGAGCTG CTCGAATTAA | 2820 |
| TCTCAAGGAT GGCTTTGCTC AAGTAAAAA TGGAGAAGTT TGGCTGAGCA ATGTTTCATAT | 2880 |
| CGCGCCTTAC GAAGAGGGCA ATATCTGGAA CCAGGAACCA GAACGTCGTC GTAAACTCCT | 2940 |
| GCTCCATAAA AAGCAAATTC AAAAATTGGA ACAAGAGATC AAAGGGACAG GAATGACCTT | 3000 |
| AGTTCCCTT AAGGTCTATA TAAAAGATGG CTACGCTAAG CTTCTTTTAG GACTTGCCAA | 3060 |
| AGGGAAGCAT GACTATGACA AACGGGAGTC TATCAAACGT CGTGAGCAAA ATCGAGATAT | 3120 |
| CGCGCGTGTG ATGAAAGCTG TTAATCAGCG ATAAAAAGAG GAATTGAAAA TGGAAAAATT | 3180 |
| AGTTGCCTAT AAACGCATGC CTTTGTGGAA TAAACAAACA ATGCCTGAAG CTGTTTCAGCA | 3240 |

796

| | |
|--|------|
| AAAGCACAAT ACAAAGTTG GGACTTGGGG GAAATTTACT GTCTTGAAGG GAGCTCTCAA | 3300 |
| GTTTATTGAA TTGACAGAAG AAGGGGAAGT TCTAGCTGAA CACCTCTTTG AAGCAGGGGC | 3360 |
| AGACAATCCA ATGGCCCAAC CTCAAGCCTG GCACCGAGTG GAAGCTGCCA CAGATGATGT | 3420 |
| GGAATGGTAC TTGGAATTTT ATTGTAAACC TGAGGATTAT TTTGCTAAAA AATACAATAC | 3480 |
| CAATCCTGTT CATTCAGAGG TCCTAGAGGC CATGCAGACA GTGAAACAAG GGAAAGCTTT | 3540 |
| GGATTTGGGT TGTGGTCAGG GGCCTAATTC TCTTTTCTA GCCCAGCAAG ATTTTGATGT | 3600 |
| GACGGCTGTA GATCAAAATG GACTAGCTCT TGAAATCTTG CAAAGCATTG TGGAGCAGGA | 3660 |
| AGATTTGAC ATGCCTGTTG GCCTTTACGA TATCAATTCA GCTAGCATTG AACAAGAATA | 3720 |
| TGATTTTATC GTTTCACAG TTGTCTCAT GTTCTACAA GCGGACCGCA TTCCAGCTAT | 3780 |
| TATTCAAAAT ATGCAGGAGA AAACCAAGTGT TGGTGGTTAC AACCTTATCG TTTGTGCCAT | 3840 |
| GGACACGGAG GATTATCCTT GCTCGGTTAA CTTCCCATTC ACCTTTAAAG AAGGAGAACT | 3900 |
| GGCAGACTAT TACAAGGATT GGGAATTGGT TAAGTACAAT GAAAATCCAG GCCATTTGCA | 3960 |
| CCGTCGCGAT GAGAATGGCA ATCGTATTCA ACTACGCTTT GCGACCTTAC TAGCTAAGAA | 4020 |
| AATCAAGTAA ACACACATGA AGATTAGGAA TTTTCTGAT CTTTTTCTT TTTTACGAAT | 4080 |
| GATATAGAAA AGGAGGGAAT TCATGTTTGT TGCAGAGAT GCTAGGGGAG AATTGGTAAA | 4140 |
| TGTGTTAGAG GATAAACTTG AGAAGCAAGC ATACACCTGC CCAGCTTGTG GAGGCCAGCT | 4200 |
| CCATTTGCGT CAAGGACCAA GTGTACGGAC GCATTTTGCC CATAAATCCT TAAAAGACTG | 4260 |
| TGATTTTTC TTTGAAAATG AAAGTCCAGA ACACCTGGCC AATAAGGAAT CCCTCTATCA | 4320 |
| CTGGTTGAAA AAAGAGACAA AGGTTCAATT AGAGTACCCG CTTTCAGAAC TTAAACAGAT | 4380 |
| TGCGGATGTA TTTGTAAATG GCAATCTAGC TCTAGAAGTT CAGTGTAGTC CCTTGCCTCA | 4440 |
| GAAAGTCCTT AAAGAGCGAA GTGAGGGCTA TCGTAGTCAG GGTTACCAAG TACTGTGGTT | 4500 |
| GCTGGGTCAA AAAGTGTCG TCAAGGAGCG TTTGACTCGT CTACAGCAAG GTTTTCTTTA | 4560 |
| TTTCAGTCAA AACATGGGCT TTTATGTTG GGAATTAGAC AAGGAAAAAC AAGTTTAAAG | 4620 |
| ACTCAAATAC CTGATTTACC AGGATCTCCG CGGTAAATC CATTATCAAA TCAAGGAATT | 4680 |
| TTCTATGGT CAAGGTAGTT TATTGGAAAT ATTGCGTCTT CCCTATAAGA GACAAAAAAT | 4740 |
| ATCTCATTTT ACAGTTTCTG AGGACAAGGA CATCTGTCGC TATATCCGGC AACAACTTTA | 4800 |
| TTATCAAAAT CTCTTTTGA TGAAAGAACA AGCAGAAGCC TATCAAAAGG GAGAAAATAT | 4860 |
| CCTGACTTAT GGAAGTAAAG AATGGTATCC ACAAATTCGA CCAATAGTGG GCAAATTTT | 4920 |
| CCGATTGAA CAAGACTTGA CTAGCTATTA TCAGCACTTT TATACCTATT ACCAAAAAAT | 4980 |
| TCCTCAAAAT GATTGGCAA AGCTTTATCC ACCAGCCTTT TATCAGCAAT ATTTCTTGAA | 5040 |

797

| | |
|--|------|
| AAATATGGTA GAATAGAAAG GATGGAGGAA TCTAATGGTA TTACAAAGAA ATGAAATAAA | 5100 |
| TGAAAAAGAT ACATGGGATC TATCAACGAT CTACCCAACT GACCAGGCTT GGGAAGAAGC | 5160 |
| CTTAAAGAT TTAACAGAAC AATTGGAGAC AGTAGCCCAG TATGAAGGCC ATCTCTTGGA | 5220 |
| TAGTGCGGAT AACCTACTAG AAATCACTGA ATTTCTCTT GAAATGGAAC GCCAGATAGA | 5280 |
| GAAGCTTTAC GCTTATGCTC ATATGAAGAA TGACCAGGAT ACACGTGAAG CTAAGTATCA | 5340 |
| AGAGTACTAT GCCAAGGCCA TGACACTCTA CAGCCAGTTA GACCAAGCCT TTTTATTCTA | 5400 |
| TGAGCCTGAA TTTATGGAGA TTAGCGAAAA GCAGTATGCT GACTTTTITAG AAGCTCAACC | 5460 |
| AAAGCTGCAG GTTTATCAAC ACTATTTTGA CAAGCTTTTG CAAGGCAAGG ATCACGTTCT | 5520 |
| TTCAACAGT GAAGAAGAAT TATTGGCTGG AGCTGGAGAA ATCTTTGGTT CAGCAAGTGA | 5580 |
| AACCTTCGCT ATCTTGACA ATGCGGATAT TGTGTTCCCT TATGTCCTAG ACGATGATGG | 5640 |
| TAAAGAAAGT CAGCTATCTC ATGGGACTTA CACACGTTTG ATGGAGTCTA AAAACGTGA | 5700 |
| GGTTCGCCGT GGTGCCTATC AAGCTCTTTA TGCGACTTAC GAACAATTCC AACACACCTA | 5760 |
| TGCCAAAACC TTGCAAACCA ATGTTAAGGT GCAAAATTAC CGTGCTAAAG TTCGTAACCTA | 5820 |
| CAAGAGTGCT CGTCATGCAG CCCTCGCAGC GAATTTTGT CCAGAAAGTG TTTATGACAA | 5880 |
| TTTGGTAGCA GCAGTTCGCA AGCATTGTC ACTCTTACAT CGCTATCTTG AGCTTCGTTT | 5940 |
| AAAAATCTTG GGGATTTCAG ATCTCAAGAT GTACGATGTC TACACACCGC TTTTATCTGT | 6000 |
| TGAATACAGT TTTACCTACC AAGAAGCCTT GAAAAAGCA GAAGATGCTT TGGCAGTCTT | 6060 |
| GGGTGAGGAT TACTTGAGCC GTGTTAAACG TGCCTTCAGC GAGCGTTGGA TTGATGTTTA | 6120 |
| CGAAAATCAA GGCAAGCGTT CAGGTGCCTA CTCTGGTGGT TCTTATGATA CCAATGCCTT | 6180 |
| TATGCTTCTC AACTGGCAAG ACAATCTGGA CAATCTCTT ACTCTGTTC ATGAAACAGG | 6240 |
| TCACAGTATG CATTCAAGCT ATACTCGTGA AACTCAGCCT TATGTTTACG GGGATTACTC | 6300 |
| TATCTTTTGG GCTGAGATTG CCTCAACTAC CAATGAAAAT ATCTTGACGG AGAAATTATT | 6360 |
| GGAAGAAGTG GAAGACGACG CAACACGCCT TGCTATTCTC AATAACTTCC TAGATGGTTT | 6420 |
| CCGTGGAACA GTTTTCCGCC AAACTCAATT TGCTGAGTTT GAACACGCCA TTCACCAAGC | 6480 |
| AGATCAAAAT GGGGAGGTCT TGACAAGCGA TTTCTTAAAT AAACTCTACG CAGACTTGAA | 6540 |
| CCAAGAGTAT TATGGTTTGA GTAAGGAAGA CAATCCTGAA ATCCAATACG AGTGGGCTCG | 6600 |
| CATTCCACAC TTCTACTATA ACTACTATGT ATATCAATAT TCAACTGGCT TTGCGGCCGC | 6660 |
| CTCAGCCTTG GCTGAAAAAA TTGTCCATGG TAGTCAAGAA GACCGTGACC GCTATATCGA | 6720 |
| CTACCTCAAG GCAGGTAAGT CGGACTATCC ACTTAATGTC ATGAGAAAAG CTGGTGTGTA | 6780 |

798

| | |
|---|------|
| TATGGAGAAG GAAGACTACC TCAACGATGC CTTTGCAGTC TTTGAACGCC GTTTAAATGA | 6840 |
| GTTTGAAGCC CTTGTTGAAA AATTAGGATT GGCATAAAAT GGTGAATCG TATAGTAAGA | 6900 |
| ATGCTAACCA TAACATGCGT CGTCCTGTCG TCAAAGAAGA AATTGTAGAC TTGATGCGTC | 6960 |
| AGCGTCAAAA GCAGGTCACA GGTTTCTTGA AAGAATTGGA AGACTTTGCC CGCAAGGAAA | 7020 |
| ATATTCCTAT TATTCCTCAT GAAACGGTGG CTTATTTCCG TTTTCTTATG GAAACCATGC | 7080 |
| AGCCTAAAAA TATTCTGGAA ATTGGGACGG CTATCGGTTT TTCAGCTCTC TTGATGGCTG | 7140 |
| AACATGCGCC AAATGCTAAG ATTACAATA TTGATCGTAA TCCAGAAATG ATTGGTTTTG | 7200 |
| CCAAGGAAAA TTTTGCCAG TTTGACAGTC GCAAGCAAAT CACTCTCCTA GAGGGAGATG | 7260 |
| CGGTGGATGT CTTATCTACA CTGACAGAGT CTTATGATTT CGTCTTTATG GATTCTGCCA | 7320 |
| AGTCTAAATA CATCGTCTTT CTGCCAGAAA TCCTCAAACA TTTGGAAGTT GGTGGTGTGG | 7380 |
| TTGTCTTGA TGATATTTTT CAAGGTGGTG ATGTTGCCAA GGATATTATG GAAGTCCGTC | 7440 |
| GTGGTCAGCG AACCATTTAT CGAGGCCTTC AAAAATTATT TGATGCAACC TTAGACAATC | 7500 |
| CAGAACTCAC CGCAACATTA GTGCCTTTAG GAGATGGTAT TCTCATGCTT CGTAAAAATG | 7560 |
| TAGCAGATGT TCAACTGTCT GAAAGCGAAT GATTTCAGA AAAATTTAAG AAAAAATAGT | 7620 |
| AAAATAGATA GAGTAACACT TATCTCAAAG GAGTAGACAT GAAGAAAAAA TTATTGGCAG | 7680 |
| GTGCCATCAC ACTATTATCA GTAGCAACTT TAGCAGCTTG TTCGAAAGGG TCAGAAGGTG | 7740 |
| CAGACCTTAT CAGCATGAAA GGGGATGTCA TTACAGAACA TCAATTTTAT GAGCAAGTGA | 7800 |
| AAAGCAACCC TTCAGCCCAA CAAGTCTTGT TAAATATGAC CATCCAAAAA GTTTTTGAAA | 7860 |
| AACAATATGG CTCAGAGCTT GATGATAAAG AGGTTGATGA TACTATTGCC GAAGAAAAAA | 7920 |
| AACAATATGG CGAAAACTAC CAACGTGTCT TGTCACAAGC AGGTATGACT CTTGAAACAC | 7980 |
| GTAAAGCTCA AATTCGTACA AGTAAATTAG TTGAGTTGGC AGTTAAGAAG GTAGCAGAAG | 8040 |
| CTGAATTGAC AGATGAAGCC TATAAGAAAG CCTTTGATGA GTACACTCCA GATGTAACGG | 8100 |
| CTCAAATCAT CCGTCTTAAT AATGAAGATA AGGCCAAAGA AGTTCTCGAA AAAGCCAAGG | 8160 |
| CAGAAGGTGC TGATTTTGCT CAATTAGCCA AAGATAATTC AACTGATGAA AAAACAAAAG | 8220 |
| AAAATGGTGG AGAAATTACC TTTGATTCTG CTTCAACAGA AGTACCTGAG CAAGTCAAAA | 8280 |
| AAGCCGCTTT CGCTTTAGAT GTGGATGGTG TTTCTGATGT GATTACAGCA ACTGGCACAC | 8340 |
| AAGCCTACAG TAGCCAATAT TACATTGTAA AACTCACTAA GAAAACAGAA AAATCATCTA | 8400 |
| ATATTGATGA CTACAAAGAA AAATTAATAA CTGTTATCTT GACTCAAAAA CAAATGATT | 8460 |
| CAACATTTGT TCAAAGCATT ATCGGAAAAG AATTGCAAGC AGCCAATATC AAGGTTAAGG | 8520 |
| ACCAAGCCTT CCAAATATC TTTACCCAAT ATATCGGTGG TGGAGATTCA AGCTCAAGCA | 8580 |

799

| | |
|--|-------|
| GTAGTACATC AAACGAATAG TCCAAATCAA TGAGTCAGGG AAAAACTCG ACTTCAGGAA | 8640 |
| AAAATGAAGC AAACATTCCC ACAATAAAAC GCATAGTACA AGGTTTGTAC TGCCCCCAA | 8700 |
| AAAGTTAGAC AATTAATTTA TCCGAAGGAT TTAGTTCTGT ATTGCACAGA GCTAAGTCCT | 8760 |
| TTTAGTTTTA TCTTAATTCT CTTATTGTTG TAATAATCAA TATAGTCTAT AATGGCTCGT | 8820 |
| TCCAATTGAT TAAGTGATTT AAATGTTTTC TCATAGCCAT AAAACATTTC GGATTTTAAA | 8880 |
| ATGCCAAAGA AAGATTCCAT CCTACCGTTG TCTTGGCTGT TGCCCTTACG TGACATGGAT | 8940 |
| GCTTGAATTC CCTTACTCTC TAGGAAGCGA TGATAAGAAT CGTGTGATA TTGCCAGCCT | 9000 |
| TGGTCACTAT GGAGAATCGT ATTCTCGTAG TGCTTCTCTT TGAATGCCTG TTCCAACATT | 9060 |
| AACGATCAAT CAATTTAATC ATGTACCTAA GATTAGAATT GTTTATCCCA AATTTATTTG | 9120 |
| AAAGCTTCTC TAAGCTATAT CCTTGTTTTC TAAGTTCATA GATCTGAACT TTATCATCAT | 9180 |
| AAGTTAATTT CATAATAAAA ACACCCCAA AGTTAGATTT TTTCTGTCTA ACTTTTGGGG | 9240 |
| TGTAGTTCAT GTACACCTGA TATGATGCGT TTTATAATTT TAAAGACTTT TTGACCAGCC | 9300 |
| TCATTTTMTT AACTTGATAC TCAGTGAAAA GCAAAGATTA AACTAGGAAG CTAGCTGTAG | 9360 |
| GCTGCTCAAA GAACAGCTTT GAGGTTGTAG ATAAACTTG TGAGGTCACC AACATATATA | 9420 |
| ATGTGAAGCT GACGTGGTTT GAATAGATTT TAGAAGAGTA TGAGTCTGGA AGTTTAAATG | 9480 |
| GATAATGCAA GATTCCATAG AATGGGTAAG CTAGAGTTCT TATGTGAAGA GTTTGGGCAT | 9540 |
| AACTTTTAC CTTTTCCTCC CTACTCATCT TAGTATAGAA AAGTGAATCT GAAATAGTAC | 9600 |
| ATAACTGCTT CTAACAATTT CTTATAAATT GATTTAAATT CTCAAATCAT ATTATTCAGT | 9660 |
| TCTTATTTCA TTTTGTCTA CAATCCTGTT GAGAAGACAC GTGTTTCATAT CAAAAGGTA | 9720 |
| TTGGCAAGTT GCAATACCTT TTTACGAGGC TCTGTTGTCT TATTTTGTGTT TCAACTGACT | 9780 |
| ATATCTCCTA TGGTTCAGT TCAGAAGGCT AGGCTATAAT TATGATTGAT AAGAAGTATC | 9840 |
| ATTCCAAGTA TTGGGAGTGA ATGTTTCAAA ATCATGGGTT TCTATAATGG TCAGGCTGGC | 9900 |
| ATTTGCTAGA CCGCCATCTT TACGAAGAAG TGGTTCCTTA TAGCCTAGGA GAGTACGAAG | 9960 |
| ACTGGCAGTA AGATTGGCGC CGTGTCGAC AATTAGAATA CGTTCAGCTG GACTATCTTT | 10020 |
| TAATGATTTG ATAAATGGGA TGGTCCGTTG AGTTGTAATA TAGAGGGATT CGGCTCCGAA | 10080 |
| CATTCGAGTG TCAAATGAG CAAGATTGA ACGAAAAGCC TGGATTGTT GCGGGTAAAT | 10140 |
| AGCTTCCAAG GTTGCAATTT TCAAACCTTC TAACTTCCA AGTTGCCATT CACGGAGATT | 10200 |
| AGGAACGATT TCTAAAGAAC AGGGGTATA GAGTTGACTT TGGATAATCT CAGCAGATTT | 10260 |
| GACCGCTCGA GGTAATCAC TTGAATAAAT CTGATCAAAA GGAATTTCCT TGAGATACTG | 10320 |

800

| | |
|--|-------|
| ACCAAGTCGT TTTAGGGTTT CAATGGATTC AGGAAGAAGA GGAGAATCAC CACTAGCACC | 10380 |
| TTGAAAACGA CCTTCTTGGT TCCAGAGGGT ACGACCCTGG CGGACAAAGT AGAGTTTCAT | 10440 |
| TACTTGATGT CCTCCAAAAT ATCTACAAAG TCTGCCTTTA CAAAGCTAGC CAAGTCTTGT | 10500 |
| GGCGCGACGA TAATGCTGTG TCCGACTTCG CCTGCAGAGA CAATCATTTG ATCCAAATCT | 10560 |
| AGAGCAATTT TATCGATAAA AATGGGATAA TTGTGTTTCT GACGAATTCC GACAGGATTA | 10620 |
| TTGGCTCCAT GAATGTAACC AGTTGTTTTT TCTAAGTCCT TTTGTGGAAT CATGCTCACT | 10680 |
| TTTTTATTGC CAGAAATTTT AGCTAGTTTC TTTTCAGACA AGTGCTGAGT GATAGGGACA | 10740 |
| ATTCGATAA TCGGTCCGGT CTTGTCTCCC AAAAGCGCCA AGGTTTTGAA AATCTGATCT | 10800 |
| CGTTCATAAC CTTGAGGAAG CTCCTCTCT AGGGCATTGA TTTGAATCCC CTGATGAGGG | 10860 |
| ATAGCTGCTT TAGATAGGAT TTGTTCCACC AATGTTTTT TGATTTTAAC TTTTTTTGCC | 10920 |
| ATTATTTATA TTTATCCTCC AATTGACTCA TCCAAATACC AAGCCAGATT CCCAGCGCAA | 10980 |
| AGAAGAAGGC GATGATGACA TAACCGACAA GTGAAAGTCC TGTGTATTGG ATACTTTTCAG | 11040 |
| CGTTTCCTGC ATTTGGAATT AAGATCAAAA GGGTACTTGA TAGGACGATA CCGATGATGA | 11100 |
| AATGATAGAC GAACTGTTTA CGGAGTTCTT CTAGTTCTCC GTCCGTCCAA GCGTAGGCCA | 11160 |
| CTTCTTCTTT CTTGCCTTTA CCTTTGGACA TCTTGTAAG AGGTGGGAGG GCAATATAGA | 11220 |
| CATGACCTGC CTCGACTAGC GGACGCATGT AACGGTAGAA AAATGTCAAG AGCAAGGTCT | 11280 |
| GGATATGGGC ACCGTCGGTA TCCGCATCG | 11309 |

(2) INFORMATION FOR SEQ ID NO: 109:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 5548 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 109:

| | |
|---|-----|
| CCATAGTCTA ACAAGTCTTT GTAAAGGTTT ATCCCTGATT CATGTAAAGA TTGTGTAAAG | 60 |
| AATCAAAAAA AGCCACTTTT GAAAAATGGC TGCTCCTAAA AATAGCTTTA AAAATTATTA | 120 |
| GTCCTGTGCG AAAGATTGGT TAGGAAGAAA AATCGTGAAG CAACTGCCTC TGCCAAGCTG | 180 |
| ACTCGTCACC GTGACTTGGC CACCTAATAA TTGACTGAGT TCTTTGACAA TGGCAAGGCC | 240 |
| AAGACCAGTG CCACCAGTTT GTCTGCTTCG ACCTTTATTA ACTCGGTAAA AACGTTCAAA | 300 |
| AATACGATCC TGCTCTAATT GACTAATACC AATCCCTGTA TCTGATACAG AAATCTTAAT | 360 |
| GCCTTCGTTT ACCTTTTGGG TCTTGACCTC AATTTTCCC CCTTGTTTCTG TGTAACGGAT | 420 |

801

| | |
|--|------|
| GGCATTGGAT AAAAGATTGA GTAAGATTGT GGAAAGTAAT TGACTATCTG ATACGAGGGT | 480 |
| GACATCATCT GGCACCTGCA CCTTTAGCTG TAAATCCTTC TTCTTGAGCT GAGGTTGCAA | 540 |
| GCTTTGAGTC AAATCCTGTA CAAATTCTGC CAAAGAAAGG GTCGTCCATT GTATAGGCAT | 600 |
| TTGTTGAGCC TTAGATAAGG TAAGAAGATG CTCAACAATA TGCTCAAGAC GCAAACCTTC | 660 |
| TTTGTAATA ATGTCTAGAA AGTCATCCTT GAGCGCTTCT TCTTCAGCTG ACATCCCCTT | 720 |
| AATGGTTTCA GCAAAGCCCT TAATCGAAGT AACTGGTGTC CTCAATTCAT GGGAGGCATT | 780 |
| TGAGACAAAG GCTAAATTTA ACTTTTCATA AGTTCTAATC GTTGTAAAT CATATAGCAA | 840 |
| GACGAGCACA GCTTCCACAG ATTGGGTGGG GCTAAAAACG GGAAGTGTG TCACCTCTAA | 900 |
| AATCAAGTCA CCCTCATGAA ACCCACTTAC TTCTTGTTTT AACCTTGTTT TTTGATCAAA | 960 |
| GGCTTGGTGA ACTAAATTCC GAATATCCAT CCGTTTGAGG TCATCAAGTG AACTTATGTC | 1020 |
| GCCGTCCACA TCGGGAAAAT AATGAGGCAG AGAGCGACTG GATAATAACA TCTGACCTTG | 1080 |
| AGCGGAAACT AAAAACGTCC CCATGGTTAG GTGCGACAGA AGAACCTCCA TTGTTTCGGC | 1140 |
| TAGATCCTTG TATTGCTGAT CCTGTTGGGA GACTTTGGTT TTTAGGCCAG ACACATACTG | 1200 |
| AGCCAAAGAC TTTAAGTCTT CTTGCCCTTT TTCTAAAAAG TATTCACTAC TGGTCAAGAG | 1260 |
| AGGTGTTGTC AAGGTCTCAA AAGCAACTTC CCATTTCCAA AGGCAAAAGA GCCAGTAGCC | 1320 |
| ACCTAGTCCC AAAGAAAGGG CTAGAAGAAA GAGACCGATG CCTTTACTGA TCCAAGTTAA | 1380 |
| TGCCATCCCT GCAATCAGAA TGAGGCTAAC ACTTAGATTG ACTAGCCAAA ATTGAAGGTA | 1440 |
| GCGTTTCATC TATAACTCCT TGAAGTTATA ACCATAACCC CGAATGGTTC GAATAAATTG | 1500 |
| AGGGGCTTTA GGATTGTCTT CAATTTTTTC CCTCAACTTA CCAATATGAA CGTCCACCAA | 1560 |
| ACGTGTTTCC TGCCCAAAGT CATACCCCA GATACGTTCC AAAAGACGCT CTCTAGTCAG | 1620 |
| TGTCATGTTG GGATGTTTCA TAAGATAGAG CAAGAGTTCA AATTCCTTTG GGGTCAAAC | 1680 |
| CAGTAACTTA TTCGCCTTGT AGACTTCATG ACGCTCAGGG TATACTTTCA AGGTCCCAAA | 1740 |
| TAGCCAAGAA TCGTCAGCGA TATTATCTGA ATCATCTCCT TCTTGTTCTC CTTTAGTTTCG | 1800 |
| CCTGAGGACA GCCTTGACAC GCGCCAGCAA TTCTCTAGGG CTAAAAGGCT TGGTCAGGTA | 1860 |
| GTCATCAGCC CCTAATTCCA AGGCCAAAAC CTTATCAAAT TCATCACTTT TCGCAGAAAC | 1920 |
| CATCATAATT GGAGTTTGA CGCCTTTGGC TCTCAGCCGC TTACAAACTT CCATGCCATC | 1980 |
| TAATTGTGGT AACATGATAT CAAGCAAGAT AAAATCAAAG GGTTCGTGTT CTGCCAAAGC | 2040 |
| TAAGGCCTTC CGTCCATTTG TCACCAATTG AGTAGAAAAG CCTTCCTTAC TTAAATGGTA | 2100 |
| GTCAAGCAAT TTCAGAATGT GTTCTTCATC ATCCACTAAT AAGACTTGTT TTGTCATCTA | 2160 |

802

| | |
|---|------|
| TTATCTCCTA TTGGTAACAT TATAACACAA TTATCAGAAA TCCTAACATT GCTAAATCAG | 2220 |
| ATTAAATTTG CCTATCAAGA CTAGTATCTG GTCAAACGCT CAATCATCTC CTTGTGCTCT | 2280 |
| GGATAGGTCG CCAGTAGATC TACCCTTTCA AATAATTCAA AATCCTCAAA TTCAAAACCA | 2340 |
| GGAGCAACAA GACAAGAAAC CAGAGCATCA TCCTTATCAA CTGTGATCC CCAAATAGTG | 2400 |
| CCCTTAGGAA CACAGTAGTG AAGTTGTTGC CCTTTGGATA TGTCCAGGCC TAAAGTGA | 2460 |
| GCTTCGTAGT GACCATCTGC TGTAATCATG TGAACAGTAA GTGGGGATCC TGCATGAAAA | 2520 |
| TACCAGATTT CATCTGCTGT CAATCGGTGA AAATGTGAAG GATTCGTTTC TTCTAATAAG | 2580 |
| AAATAAATAC TGGTATAAAG CGCCCTTCCC TTACCAGCAA GGTTTATAGT GTCTGAAGCT | 2640 |
| TTTTTTGTTT GTCTAAAATA GCCACCTTCA ATATGGGGAG CTAAGTCTAG AGTTCTTATC | 2700 |
| AAGTCTTCTT TATCCGTCGG AGCCAATGGG TTGAAGTAAC TCTTGTTCAA AGTGGTTTTA | 2760 |
| CGATTTCAAG AACTCCTCTC AGTTCTGAGG ACACGGTAAT GATTGATGCG ACGGAAGTAC | 2820 |
| AAATCAATCG CCCTAAAAAA AGAATTAGCG AATGATTCTG GTAAAAAAA TGCCACGCTA | 2880 |
| TGAAGGCTCA AGCGATTGTC ACAAGTCAAG GGAGAATTGT TTCTTTGGAT ATCGCTGTGA | 2940 |
| ACTATTGTCA TGATATGAAG TTGTTCAAAA TGAGTCGCAG AAATATCGGA CAAGCTGGTA | 3000 |
| AAATCTTGGC TGACAGTGGT TATCAAGGGC TCATGAAGAT ATATCCTCAA GCACAACTC | 3060 |
| CACGTAAATC CAGCAAACTC AAGCCACTAA CAGTTGAAGA TAAAGCCTAT AACCATGCGC | 3120 |
| TATCCAAGGA GAGAAGCAAG GTTGAGAACA TCTTTGCCAA AGTAAAAACG TTTAAAATGA | 3180 |
| TTTCAACAAC CTATCGAAAT CATCGTAAAC ACTTCGGATT ACGAATGAAT TTGATTGCTG | 3240 |
| GCATTATCAA TCATGAACTA GGATTCTAGT TTTGCAGGAA GTCTATTATT TGGTTAGGTG | 3300 |
| AATTAGTGAA GCGTTTGGC AAGTGTCTCT GGTTACGACG TCATGGACTC TAAATCGATT | 3360 |
| ATATTTAGGG GTCATGACTA GTGAAGCAGT TAGCTAGTTC GCATATAAGC GGCTAGCGTC | 3420 |
| TAACAATTAG GAACTTTAGT TCCAATAACT TTAAGATTAC GACGTTTTAG GACATAAATC | 3480 |
| GATCATATTT ATGTCCTAAA ACTAGTGAAG CGCCTAGCCA AAGTCCGAAT AGGATTTGGC | 3540 |
| GTAGTTACT TAGATTGCTT TGCAATCAAG TAACTTTGGC GATTTACATC TTCTCTGGCG | 3600 |
| CTTCTACTCC AAGCAAGCGA AGGGCTTCTT TGAGAACGAC TCGGTTGCG TAGCTGAGGG | 3660 |
| CTAGACGGCT GTCGCGTTCT GGGCTTTCAT CCAAGATACG TGTATGTGCA TAGTATTTGT | 3720 |
| TAAAGGATTG AGCCAGGCTA ATTGCAAAT TAGCAATGAT AGAAGGTTCA AAGTTATCTG | 3780 |
| CCGCACGGTT GATAATACGT GGGAAGTCTT GAATGAGTTT AATGATTTCC CAGCTTTCAG | 3840 |
| TATCATTCAA GCTATAGTTG CCAGCTGTTT CTGGTTTGAA ATCGGCTTTG CGTAAGATAG | 3900 |
| ATTGGATACG AGCGTAGGCA TATTGAACGT AAGGTCCAGT TTCACCTCG AAGGATACCA | 3960 |

803

| | |
|--|------|
| TAGCCTCTAG GTCGAAGTCG TATCCATTG TACGGTCGGT TTTGAGGTCA TAGAATTTAA | 4020 |
| TGGCTCCAAT CCCAACAGCA TGTGCTACTT GGTCTTTGTT TTCTAGTTCA GGATTTTGTAG | 4080 |
| CCTCGATTTG GACCTTGGCA CGGCTAACAG CCTCTGCAAC AGTAGGCTCT AGCAAGATGA | 4140 |
| CATTCCCTTT ACGAGTAGAG AGTTTCTTCC CTTCCTTTGT AACCAAACCA AAAGGAACGT | 4200 |
| GAGTAATGTC GTCACTCCAG TCGTAGCCCA TCTCTTGCAA GACAGCTTTG AGCTGTTTAA | 4260 |
| AGTGGGCAGA TTGTTCTTGA CCAACGACAT AGATAGATTT AGCAAATTGG TATTTCGTTTT | 4320 |
| TACGGTAGAG GGCTGCAGCC AAGTCACGTG TGATATAGAG AGTTGCACCA TCAGACTTCT | 4380 |
| TGATGAGGGC TGGATGTTCA ATTCCATATT TCTCAAGATT CACAACCTGG GCACCTTCTG | 4440 |
| ATTCAAGAAG TAGTCCTTTT TCAGAAAGAA TGTCTACAAC TGCATCCATC TTATCATTTG | 4500 |
| AGAAGGCTTC TCCGTATAG CTGTCAAATT CAACCTTCAA TTCATTGTAA AGGCGGTTAA | 4560 |
| ATTCCACTAA ACTTTCATCG CGGAACCATT GCCAAAGAGC GAGAGCTTCC TCATCTCCAT | 4620 |
| TTTCAAGTTT ACGGAACCAT TCGCGCGCTT CTTTCATCCAA GCTAGGGTCA TTTTCAGCTT | 4680 |
| CAGCGTTGAT GCGGACATAG AGTTTAAGGA GTTCATCGAT TGGATGAGCT TTTACAGCTT | 4740 |
| CTTCGTCGCC CCATTTTTTG TAGGCAACAA TCAACATCCC AAATTGTTTA CCCCAGTCTC | 4800 |
| CCAAATGGTT GACCTTGACC GTTTGATAAC CGATTTTTTG GAAAATATGT GACAAGCTAT | 4860 |
| CTCCGATAAC AGTTGAACGC AGGTGGCCAA TAGAAAATGG TTTAGCGATA TTCGGACTAG | 4920 |
| ACATGTCGAT AACAAACATT TCTTGTTTAC CAATATTTTG GTCAGCATAG TGTCTTTTTT | 4980 |
| CAGTGGTAAC AGCTTGCAAT ACTTGAGCAG AAATGGCAGA TTTATCAAGG AAAAAGTTAA | 5040 |
| CGTAAGGTCC TGTGCGACA ACTTTTCAA AGGCTTGGCT GTTCATTTTT TCAGCCAGTT | 5100 |
| CAGCCGAAT CATTTGTGGT GCTTTACGTT CGACTTTTGC AAGAGAAAAA GCAGGGAAAG | 5160 |
| CAATGTCTCC CATTTCTGAG TTTTAGGGG TTTCCAGTAA CTTTAAAATA GCCTCTTGGT | 5220 |
| CCAGGCTATC AATGATGCTA GATAATCGC TAGCAATCAA TTCTTTTGTA TTCATTAAGA | 5280 |
| GCTCCTTTTT GGACTTTTCT ACTATTTTAT CACAATTTTA AAGAAAGAAG AAAAAATTTT | 5340 |
| TGAAATCTCC TGTTTTTTTG GTATAATATG GTTATAAATA TAGTTATAAA TATGCACGCA | 5400 |
| AGAGGATTTT ATGAGAAAAA GAGATCGTCA TCAGTTAATA AAAAAATGA TTAGTGAGGA | 5460 |
| GAAATTAAGT ACACAAAAAG AAATTCAAGA TCGGTTGGAG GCGCACAATG TTTGTGTGAC | 5520 |
| GCAGACAACC TTGTCTCGTG ATTTGCGG | 5548 |

(2) INFORMATION FOR SEQ ID NO: 110:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3132 base pairs

804

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 110:

| | |
|--|------|
| TACCCGGTAG TCTTAGCAGA CACATCTAGC TCTGAAGATG CTTTAAACAT CTCTGATAAA | 60 |
| GAAAAAGTAG CAGAAAATAA AGAGAAACAT GAAAATATCC ATAGTGCCTAT GGAAACTTCA | 120 |
| CAGGATTTTA AAGAGAAGAA AACAGCAGTC ATTAAGGAAA AAGAAGTTGT TAGTAAAAAT | 180 |
| CCTGTGATAG ACAATAACAC TAGCAATGAA GAAGCAAAAA TCAAAGAAGA AAATTCCAAT | 240 |
| AAATCCCAAG GAGATTATAC GGACTCATTT GTGAATAAAA ACACAGAAAA TCCCAAAAAA | 300 |
| GAAGATAAAG TTGTCTATAT TGCTGAATTT AAAGATAAAG AATCTGGAGA AAAAGCAATC | 360 |
| AAGGAACTAT CCAGTCTTAA GAATACAAAA GTTTTATATA CTTATGATAG AATTTTAAAC | 420 |
| GGTAGTGCCA TAGAAACAAC TCCAGATAAC TTGGACAAAA TTAAACAAAT AGAAGGTATT | 480 |
| TCATCGGTTG AAAGGGCACA AAAAGTCCAA CCCATGATGA ATCATGCCAG AAAGGAAATT | 540 |
| GGAGTTGAGG AAGCTATTGA TTACCTAAAG TCTATCAATG CTCCGTTTGG GAAAAATTTT | 600 |
| GATGGTAGAG GTATGGTCAT TTCAAATATC GATACTGGAA CAGATTATAG ACATAAGGCT | 660 |
| ATGAGAATCG ATGATGATGC CAAAGCCTCA ATGAGATTTA AAAAAGAAGA CTTAAAAGGC | 720 |
| ACTGATAAAA ATTATTGGTT GAGTGATAAA ATCCCTCATG CGTTCAATTA TTATAATGGT | 780 |
| GGCAAAATCA CTGTAGAAAA ATATGATGAT GGAAGGGATT ATTTTGACCC ACATGGGATG | 840 |
| CATATTGCAG GGATTCCTGC TGGAAATGAT ACTGAACAAG ACATCAAAAA CTTTAACGGC | 900 |
| ATAGATGGAA TTGCACCTAA TGCACAAATT TTCTCTTACA AAATGTATTG TGACGCAGGA | 960 |
| TCTGGGTTTG CGGGTGATGA AACAATGTTT CATGCTATTG AAGATTCTAT CAAACACAAC | 1020 |
| GTTGATGTTG TTTCGGTATC ATCTGGTTTT ACAGGAACAG GTCTTGTAGG TGAGAAATAT | 1080 |
| TGGCAAGCTA TTCGGGCATT AAGAAAAGCA GGCATTCCAA TGGTTGTCGC TACGGGTAAC | 1140 |
| TATGCGACTT CTGCTTCAAG TTCTTCATGG GATTTAGTAG CAAATAATCA TCTGAAATG | 1200 |
| ACCGACACTG GAAATGTAAC ACGAACTGCA GCACATGAAG ATGCGATAGC GGTCGCTTCT | 1260 |
| GCTAAAAATC AAACAGTTGA GTTTGATAAA GTTAACATAG GTGGAGAAAG TTTTAAATAC | 1320 |
| AGAAATATAG GGGCCTTTTT CGATAAGAGT AAAATCACAA CAAATGAAGA TGGAACAAAA | 1380 |
| GCTCCTAGTA AATTAAAAAT TGTATATATA GGCAAGGGGC AAGACCAAGA TTTGATAGGT | 1440 |
| TTGGATCTTA GGGGCAAAAT TGCAGTAATG GATAGAATTT ATACAAAGGA TTTAAAAAAT | 1500 |
| GCTTTTAAAA AAGCTATGGA TAAGGGTGCA CGCGCCATTA TGGTTGTAAA TACTGTAAAT | 1560 |

805

| | |
|--|------|
| TACTACAATA GAGATAATTG GACAGAGCTT CCAGCTATGG GATATGAAGC GGATGAAGGT | 1620 |
| ACTAAAAGTC AAGTGTTTTC AATTTTCAGGA GATGATGGTG TAAAGCTATG GAACATGATT | 1680 |
| AATCCTGATA AAAAACTGA AGTCAAAAGA AATAATAAAG AAGATTTTAA AGATAAATTG | 1740 |
| GAGCAATACT ATCCAATTGA TATGGAAAGT TTTAATTCCA ACAAACCGAA TGTAGGTGAC | 1800 |
| GAAAAGAGA TTGACTTTAA GTTTGCACCT GACACAGACA AAGAACTCTA TAAAGAAGAT | 1860 |
| ATCATCGTTC CAGCAGGATC TACATCTTGG GGGCCAAGAA TAGATTTACT TTTAAAACCC | 1920 |
| GATGTTTCAG CACCTGGTAA AAATATTAAA TCCACGCTTA ATGTTATTAA TGGCAAATCA | 1980 |
| ACTTATGGCT ATATGTCAGG AACTAGTATG GCGACTCCA TCGTGGCAGC TTCTACTGTT | 2040 |
| TTGATTAGAC CGAAATTAAA GGAAATGCTT GAAAGACCTG TATTGAAAAA TCTTAAGGGA | 2100 |
| GATGACAAA TAGATCTTAC AAGTCTTACA AAAATTGCCC TACAAAATAC TGCGCGACCT | 2160 |
| ATGATGGATG CAACTTCTTG GAAAGAAAA AGTCAATACT TTGCATCACC TAGACAACAG | 2220 |
| GGAGCAGGCC TAATTAATGT GGCCAATGCT TTGAGAAATG AAGTTGTAGC AACTTTCAAA | 2280 |
| AACACTGATT CTAAAGGTTT GGTAACTCA TATGGTTCCA TTTCTCTTAA AGAAATAAAA | 2340 |
| GGTGATAAAA AATACTTTAC AATCAAGCTT CACAATACAT CAAACAGACC TTTGACTTTT | 2400 |
| AAAGTTTCAG CATCAGCGAT AACTACAGAT TCTCTAAGT ACAGATTAAA ACTTGATGAA | 2460 |
| ACATATAAAG ATGAAAAATC TCCAGATGGT AAGCAAATG TTCCAGAAAT TCACCCAGAA | 2520 |
| AAAGTCAAAG GAGCAAATAT CACATTTGAG CATGATACTT TCACTATAGG CGCAAATCT | 2580 |
| AGCTTTGATT TGAATGCGGT TATAAATGTT GGAGAGGCCA AAAACAAAAA TAAATTTGTA | 2640 |
| GAATCATTTA TTCATTTTGA GTCAGTGGAA GCGATGGAAG CTCTAACTC CAGCGGGAAG | 2700 |
| AAAATAAACT TCCAACCTTC TTTGTCTGATG CCTCTAATGG GATTTGCTGG GAATTGGAAC | 2760 |
| CACGAACCAA TCCTTGATAA ATGGGCTTGG GAAGAAGGGT CAAGATCAAA AACACTGGGA | 2820 |
| GGTTATGATG ATGATGGTAA ACCGAAAATT CCAGGAACCT TAAATAAGGG AATTGGTGGA | 2880 |
| GAACATGGTA TAGATAAATT TAATCCAGCA GGAGTTATAC AAAATAGAAA AGATAAAAAAT | 2940 |
| ACAACATCCC TGGATCAAAA TCCAGAATTA TTTGCTTTCA ATAACGAAGG GATCAACGCT | 3000 |
| CCATCATCAA GTGGTTCTAA GATTGCTAAC ATTTATCCTT TAGATTCAAA TGGAAATCCT | 3060 |
| CAAGATGCTC AACTTGAAAG AGGATTAACA CCTTCTCCAC TTGTATTAAG AAGTGCAGAA | 3120 |
| GAAGGATTGA TT | 3132 |

(2) INFORMATION FOR SEQ ID NO: 111:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14672 base pairs

806

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 111:

| | |
|--|------|
| CGAGATTTCT TTAAATGAAC TACGTGAAAT CTACCCATCA TCCAGATCTG GATATTCTCT | 60 |
| CCTATCTATA AGTAAAGTTT TAGGAGATTT TAATATAAGT TCTCATGCTT TTAAAGCTTC | 120 |
| GGTAAGAGAT TTAACCCGC TCAGTTTCCC ACTCATTTGC TTCTGGGAGA GTTCTCATT | 180 |
| TATTATTCTT GAAAAATTA GTAAAAACAA GTTTTATATT TTAGATCCTG CAAAAGGCAG | 240 |
| GCAGAGAATG TCAATAAGTG AATTGAAAG GCATTATTCA AATATCATTT TAACATTAA | 300 |
| AAAGTTAGAT AGCTTTATGT CTCGTAAAGA TAATAAGAAG TCGCCTGTTT TAAAGTATTT | 360 |
| TTTTAAGTAT AGGAATAAGC TAGGGATTTT ATTTTGTGTA ACAGCATTAT TGTATGTAAT | 420 |
| ACAATCATTA GTACCTATAG CTAATAGATA CATAATTGAC ACGAATTTCAGGACGATTC | 480 |
| GTATTCGTCT AGAATGTTAT TTAATATATT ATTTATATTT ACTGTTTCAT TCTCACTAAT | 540 |
| GTATTTATTA AGACAGATAT ATGTTGCATC CTTAAATAT ATAATGGATA AAGAGATTAG | 600 |
| CTATGATTTT ATGAAACATT TGATATATTT ACCTTACAGT TTTTATGAAA AACGTACTTT | 660 |
| AGGGGATATA CTTTTAGAG CTAATCTAT TGTATATATA AGAGAAATAC TATCAAATAA | 720 |
| TTTTATAGCA GCTATACTTG ATTTGTTAAT GATTGTGGT TATGCTGTGG TTTTATTAG | 780 |
| CTTTTCTAAG TACATGGTAA TCTTTTAAAT ATCACTAAGT CTAGCTCTAT CTATTGTAAT | 840 |
| GTATCCAATC ATAAAAATCT CAAAAATTT AATTGATAAA AATATAAAG AAAAGGTAA | 900 |
| TGTTCAAAAT ATTACTTCCG AAGTAATTTC TAAAAATAGT GATATTAAGC TAACTGGAGA | 960 |
| AGAGGAATTT TGGATTAACA AATGGGATAA TTTTAATACA AAACAGCTCA TCATAGGTCTG | 1020 |
| AAAATTTGAT ATACATTTAT CAATTGTTAG TAGTATAACG AATGTTTAC AAATTATTCT | 1080 |
| CCCTGTTTGG ACCCTTATTG TAGGTGTAAA TATAAAAACA TTCGAACAAT TGACGTTAGG | 1140 |
| ACAAATTGTA GCAATAAGTA CAGTCTCACC ATACTTTATT TCTCCTATAA TTTCTTTAAG | 1200 |
| TGATAACTAT ATACAATTAA TGTATTAAA GGGATATTTT TTAAGAATAG AGGATGTGTT | 1260 |
| TAATACTAAA TCCGAATTAA TTCCAGAAAG AGTCAGTCAA GATATAAAAT TTGATAAAAA | 1320 |
| AATAGAATTA AAAGATATTT GGTATAAATA TGGATTATTT GATGATTATG TTTTGAAAGG | 1380 |
| AATAAATGTT ACTATTAAAA AAGGAGAAAC TGTGCTATT GTTGGAGAAT CAGGTTTCAGG | 1440 |
| TAAGAGTACA TTAGCTAAAA TTTTATTAGG TTTATTAGAA CCTAATATTG GTTCAATAGA | 1500 |
| AGTTGATGGA GTAGAAAAAG AAGAAATGG TCAACATTG TATAGAAAGA TTTTGGAGC | 1560 |

807

| | |
|---|------|
| AGTGTTACAA AATTCAACCC TAAGTTATGG TACCTTAAGA GAGAATTTGA CATTTGGACA | 1620 |
| CTTTGTTTCA GATGAAGAAT TAATGACAAA TCTAAATTC AATTGGTCTTA GCAATGTAGT | 1680 |
| TAAATCTTTA CCTCTTGGAT TAGAGACAAT CATCGCTGAA GAAGGTAATA ACTTTTCTGG | 1740 |
| AGGGCAGCAG CAAATGATAC TTTTAGCTCG TTGTCTTTTG TCGAAACCTT CGGTAGTTGT | 1800 |
| TTTGGACGAA GCAACAAGTA GTTTAGATAA TTTATCTCAA CAAATTACAA CTTCTTACTT | 1860 |
| AAGTGAAATC GGTACCACTA AGATTTTAAT TGCCCATCGA CTAGATACTA TCAAGTCTGC | 1920 |
| AGATAAGATC TTAGTAATGC ATAATGGTGA AATTGTAGAG ATTGGGACCC ATAGAGAACT | 1980 |
| TCTTGAACTA GGAGGCATTT ATAAGCAATT GTATTCAAAT AATTAGTTTT TGATTAAAAG | 2040 |
| GGTAAATTTA TGAAGATTAT GAAAAAAAAA TATTGGACTT TAGCGATATT ATTCTTTTGT | 2100 |
| TTGTTCAATA ATTCTGTTAC TGCTCAAGAA ATACCTAAAA ATCTTGATGG CAATATAACT | 2160 |
| CACACTCAGA CTAGCGAAAG TTTTCTGAA TCTGATGAAA AACAGGTGA CTATTCTAAT | 2220 |
| AAAAATCAAG AAGAAGTAGA CCAAAATAAA TTTCTGATTC AAATCGATAA GACAGAATTA | 2280 |
| TTTGTAACAA CAGATAAACA TTTAGAAAAA AACTGTTGTA AATTGGAACT TGAACCACAA | 2340 |
| ATAAATAACG ATATTGTTAA CTCTGAAAGT AATAATTTAC TAGGCGAAGA TAATTTAGAT | 2400 |
| AATAAAATTA AGGAAAATGT TTCTCATCTA GATAATAGAG GAGGAAATAT AGAGCATGAC | 2460 |
| AAAGATAACT TAGAATCGTC GATTGTAAGA AAATATGAAT GGGATATAGA TAAAGTTACT | 2520 |
| GGTGGAGGCG AAAGTTATAA ATTATATTCT AAAAGTAATT CTAAAGTTTC AATTGCTATT | 2580 |
| TTAGATTTCAG GAGTCGATTT ACAAATACT GGATTACTGA AAAATCTTTC AAATCACTCA | 2640 |
| AAAAACTATG TCCCAATAA AGGATATTTA GGAAAAGAGG AGGGAGAGGA AGGAATAATA | 2700 |
| TCAGATATTC AAGATAGATT AGGTCATGGT ACGGCTGTTG TAGCTCAAAT TGTAGGGGAT | 2760 |
| GACAATATTA ATGGAGTAAA TCCTCACGTT AATATTAACG TCTATAGAAT ATTTGGTAAG | 2820 |
| TCGTCAGCTA GTCCAGATTG GATTGTAAAA GCAATTTTTC ATGCTGTAGA TGATGGCAAT | 2880 |
| GATATTATCA ATCTTAGTAC TGGACAATAT TTAATGATTG ATGGAGAATA TGAGGACGGA | 2940 |
| ACAAATGATT TTGAAACATT TTTGAAGTAT AAAAAGGCTA TTGATTACGC GAATCAAAAA | 3000 |
| GGAGTAATTA TAGTAGCTGC ATTAGGGAAT GACTCCCTAA ATGTATCAAA TCAGTCAGAT | 3060 |
| TTATTGAAAC TTATTAGTTC ACGCAAAAAA GTAAGAAAAC CAGGATTAGT AGTTGATGTT | 3120 |
| CCAAGTTATT TCTCATCTAC AATTCGGTC GGAGGCATAG ATCGCTTAGG TAATTTATCA | 3180 |
| GATTTTAGCA ATAAAGGGGA TTCTGATGCA ATATATGCGC CTGCAGGCTC AACATTATCT | 3240 |
| CTTTCAGAAT TAGGACTTAA TAACTTTATT AATGCAGAAA AATATAAAGA AGATTGGATT | 3300 |

808

| | |
|--|------|
| TTTTCGGCAA CACTAGGAGG ATATACGTAT CTTTATGGAA ACTCATTTGC TGCTCCTAAA | 3360 |
| GTTTCTGGTG CGATTGCAAT GATTATTGAT AAATACAAAT TAAAAGATCA GCCCTATAAT | 3420 |
| TATATGTTTG TAAAAAATT CTGGAAGAAA CATTACCAGT AAAAAATGGT ATAAAAGTGT | 3480 |
| TAAATATACC AAACGTATTG AGATATGATT TGAATATGTT ACAATTAGAA TATAAAAATG | 3540 |
| AACAAAGTTG GGATAGTTTC ATAGATAATG TTAATTTAAT TGAGTTGGAA GAGAGAATTC | 3600 |
| AAACTACTAT TGGAATTAAA CAAATAAACA CACACAATAT TATTACTATT GCCCGAGAAG | 3660 |
| GGTACTCTCA AAATTATTTA CCTAACACTT CAGAAAATAC ATATAATTCA TTACAAGTCA | 3720 |
| GTTTAGTTGG AGTATTACTA CTTTTTATAA GTATGGTAAA TATTTTATGG GCTAAAAAAA | 3780 |
| GTAAATGAAA ATAAATTTG GAGCCCTCTG AAAAAAGTAAG TCCTACAGTT CAACTAAAAAT | 3840 |
| GAGTCAAAAG ATGAATCACC TTGATGTAGG GGAGTTTGTC TTATTGCTGC CTGAACACCT | 3900 |
| CCGTTTCAGAG GAAGAACATT ATAAATCTGT TTTTGAAGAC GACTTAACCA GTCGCATATC | 3960 |
| TAGTCAAGAT GAACGACAGC AAATGACTGC TACGGTAGGT TATTTAGAAT CAGGTCAGGA | 4020 |
| TCGTTTTGTG TATAATACGA CCCCTATTTC TTACCAGCAG TTTTGTGAAAG ATCCAATCAT | 4080 |
| CATTGTATA ACACCCCAAT CAACTGGTCC ACAGTCCATT TTGTTTTGGA TAGACGCAGT | 4140 |
| ACAGAACTAC GTTCTCTTTA ATCAATGTGC TGATGCCCAG GAGCTTATCC AGAGACAAGG | 4200 |
| CATTGAAAAAT TGGGTCTCAG AAATGCAAAC AGGTTACCAC AACTACATCA CATTATTGGA | 4260 |
| TAATATCCAG AGGGAACGTT GGGTAATGCT AGCAGGAGCT GTGCTTGGGA TTGCAACTTC | 4320 |
| AATCTTGTTG TTTAACACTA TGAATAGGCT CTACTTTGAA GAATTTAGAC GTGCCATTTT | 4380 |
| TATCAAACGC ATTGCAGGTC TCAGGTCTTT AGAAATCCAT CGCACTTATC TCTTTGTCTA | 4440 |
| ACTGGGTGTG TTTTACTGG GATTTGTTGC GAGTGTATTT CTTACGGTAG AGATAGGAGT | 4500 |
| TGCTTTCTTA GTCTTGTTAC TCTTTACTGG TCTATCTCTT TTACAGTTAC ATGTCCAAAT | 4560 |
| GCAGAAAGAA AACAAGATGT CCATGCTTGT TTTGAAGGGA GGTTAATATG ATTGAACTTA | 4620 |
| AACAGGTGAG TAAATCTTTT GGAGAACGAG AGTTATTTTC GAATCTTTCA ATGACATTTG | 4680 |
| AGGCTGGAAA AGTCTATGCC TTAATTGGTT CAAGTGGTAG CGGAAAAACA ACCTTGATGA | 4740 |
| ACATGATTGG GAAATTAGAA CCTTATGATG GGACGATTTT TTACCGAGGT AAAGACTTGG | 4800 |
| CCAATTATAA ATCAAGTGAT TTTTCCGTC ACGAATTGGG CTACCTCTTC CAGAACTTTG | 4860 |
| GCTTAATTGA AAACCAAAGT ATTGAAGAAA ACCTTAAGCT AGGTCTCAT TGGTCAAAAGT | 4920 |
| TGAGTCGGTC GGAACAGCGG TTGAGGCAGA AGCAGGCTTT AGAACAGGTC GGCCTGGTTT | 4980 |
| ATCTTGACCT AGATAAGCGC ATCTTTGAGT TATCGGCGG AGAATCGCAA CGGGTTGCCT | 5040 |
| TGGCAAAAAT TATCTTAAAG AATCCACCC T TATTCTGGC AGATGAGCCA ACAGCTTCAA | 5100 |

809

| | |
|--|------|
| TAGACCCAGC AACCTCTCAG TTGATTATGG AGATTTTGCT ATCTCTTCGA GATGATAATA | 5160 |
| GGCTAATCAT TATCGCAACA CATAATCCGG CAATTTGGGA GATGGCTGAT GAAGTGTTC A | 5220 |
| CGATGGATCA TCTGAAATAA AAATCCTTGT TTTTAATTGC ACGATGAGTT ACTGAAATAT | 5280 |
| TATCATGAAT CAAGAATTGG AGTTAATTGA GAATTGTAAT TAATTTAGAA TTGTACTTTA | 5340 |
| TTAATATTGA GGTAACCTTT TCTTGATAAA GGAAGAAATA ATGGAGAGGA AGTTAGAATG | 5400 |
| AAAAAATTCG ACAATTATAT TATTGAGAAG CCTTGCGATT CTAATTCAGA TAAACTGCAA | 5460 |
| AAAATCTTAA TAAATGAAAG TTTGGTAGAT GATATTTTGC AATTTTCTCT CAGAATCAAT | 5520 |
| AATAGTGTAG GAGAGATTTT CCTCCTACAA CCGTTTTAAA AGAAAACTAT CTTTATTCCA | 5580 |
| TGTTATTTTG AGGAAGATAT TGTGAAAGTC AAAGATGATG ATAAAGTTGA GTGGAATTTG | 5640 |
| TTAGAATTTT AAAAATTTAG AGCATTTTTG GCTTAGTAAT CTGTGTTGAA GGCTCAAAAC | 5700 |
| CTATGGTAAA AAAGTAGCTT TGAACGTA TTGCCTCCAA AGATTTAGTT AAATAATGAT | 5760 |
| TTAACACAAA AAGAAATTAT TGAAGTCTG GAAAGATGTT GTTTCAGTAT TGAGAAAAGG | 5820 |
| TGGGAAAAAC TTGCGATTTT CACAGAGAAA GGAAGAAAAA GTATAGAAAT ATAGTCAATT | 5880 |
| GAAACAAGAA CAGGATAAAA GAACCTTTTG TGCCATATTT TTCTCCTTTC GCTTTACAAT | 5940 |
| TGGATTGAAC ACCTTTATTT TATCGCGTTT GGAGTTTTTT TGGTATAACC TTCGACGCAC | 6000 |
| ACCCGCATAG CGGGTGTTTT TTTTGTCTCG CACCTAACGG AGCGAGACAA ACTAATAGTC | 6060 |
| ACTTAATCAA AAAACGCACC ATATCAAAAA CTAAGAGTT TGATATCATG CGTCATGTCT | 6120 |
| TAAACTAATT GACTATACTT TCTATTCAAA TGAGCTTTTA ACCAATTGAT TGAGCCAATC | 6180 |
| CACTCTTAAA ACCAAAGAGC AATTTCTCGC TTAGCTGACT CTTCTGAATC TGAACCATGT | 6240 |
| ACAACATTTT GGATAATCTC ATTTTCTCCA GCAGCTTTTG CAAAATCACC TCGAATAGTG | 6300 |
| CCTGGTAAAG CTTCTTCTGG ACGAGTTGCA CCCATCATGG TCCGCCAAGT TTCGATTACT | 6360 |
| TTGGGACCAG AAATGACACC CACAAGAACT GGACCTGAAG TCATGAATTC ACGAATCGGT | 6420 |
| GGGTAAAAAC TCTGACCAAC CAAGTCCTGA TAGTGCTGGT CAATCAACTC TTCTGAAACC | 6480 |
| TGTGAACGAA ACTCCAATTT TTCGATTGTA AATCCACGTT GTTCGATGCG CTTTAACACT | 6540 |
| TCACCCACTA GCCCTCTTTT TACACCATCT GGTTCGATGA TAAAGAATGT TTGTTCCATA | 6600 |
| CCCGTCTCCT TTGTCAGCTT CTTTCTTTTA TTTTACCACA TTTCGTGGAA AAATGGAGAA | 6660 |
| AGTTTTTCAGA AGAGAGAATG AGAGAACCCT CGGGTTCTCT CATTTCTCTT TATTCTACTG | 6720 |
| TTTCTTCCAC AGTTTCAACG GCAGTATCCA CAACTACTTC TGTGTTTCT TCATTTCTCT | 6780 |
| CTTCCTCTAC TGGAGGATTA AGGTATTCTT CTTGTTGAC AGCATGTGGT TCAAGGTTAC | 6840 |

810

| | | | | | | |
|------------|-------------|------------|------------|------------|------------|------|
| GGTAACGGGC | CATACCAGTA | CCAGCTGGGA | TGATCTTACC | GATGATAACA | TTTCTTTAA | 6900 |
| GTCCAAGGAG | ATGGTCTTTC | TTACCACGGA | TAGCTGCGTC | AGTAAGGACA | CGAGTTGTTT | 6960 |
| CCTGGAAGGA | AGCCGCTGAC | AAGAACTGT | TTGTTTCAAG | TGAGGCTTTG | GTAATTCCCA | 7020 |
| TAAGGACTGG | GCGACCTGTC | GCTGGAACTC | CACCTGCGAT | AAGGACATCT | TTGTTGGCAT | 7080 |
| CTGTAAAGTC | ATTGATATCC | ATGAGGGTAC | CCATGAGAAG | ATCTGTATCA | CCTGGATCCA | 7140 |
| TGACACGGAC | TTTACGGATC | ATTTGACGAA | CCATTACCTC | GATGTGTTTG | TCACCGATT | 7200 |
| CTACCCCTTG | GCTACGGTAA | ACTTTTGTGA | CTTCACCGAG | AAGGTACGTT | TCAACTGACA | 7260 |
| AGACATCACG | AACTGCAAGG | AGACGTTTGT | GTTGGATAGA | ACCTTCTGTC | AGAGCAGCAC | 7320 |
| CACGCGCTAC | TTGGCCCCCA | ACTTCGACAC | GCATACGAGC | TGTAAATGGA | ACGACATATT | 7380 |
| CACCTTCGCC | AGTTTCACCC | TTAACAAAGA | CTTCTTGGT | ACGAGTTGAT | GCATCTTCTT | 7440 |
| CGATAGCAGT | AACTTGTCTT | TTAACCTCTG | TAATAACCGC | TTCCCTTTTA | GGATTGCGGG | 7500 |
| CTTCAAAGAT | TTCTTGGAACA | CGAGGAAGAC | CCTGAGTGAT | ATCGGTATTT | GAGGCAACCC | 7560 |
| CACCTGTGTG | GAAGGTACGC | ATTGTAAGCT | GTGTACCAGG | TTCCCCGATA | GATTGGGCAG | 7620 |
| CGATTGTACC | AACTGCTTCA | CCAACTTCAA | CCGCATCACC | AGTCGCCAAG | TTGATACCGT | 7680 |
| AACAGTGACG | GCAGACACCG | TGACGAGTGT | TACATGTAAA | TACAGAACGG | ATAGTCACTT | 7740 |
| CTTCCACACC | AGCATTGACA | ATTTCACGCG | CCTTGTCTTC | TGTAATCAAT | TCATTTGGAC | 7800 |
| CAATAATCAC | TGCACCAGTT | TCTGGATGTT | TAACAGTTTT | CTTAGTGTA | CGACCGTTGA | 7860 |
| GACGCTCTTC | GAGAGACTCG | ATCATCTCTT | TTCCTTCTGC | GATAGAACGG | ATCAAGAGAC | 7920 |
| CACGGTCAGT | TCCACAGTCG | TCCTCACGGA | TGATAACGTC | TTGGGCAACG | TCGACCAAAC | 7980 |
| GACGAGTCAA | GTAACCTGAG | TCGGCTGTCT | TAAGGGCCGT | ATCGGTCATA | CCTTTACGAG | 8040 |
| CACCGTGAGT | TGAGAAGAAC | ATTTCCAATA | CCGACAAACC | TTGCGGAAG | TTTGAAAGGA | 8100 |
| TTGGCAATTC | CATGATACGT | CCATTGCGAG | CAGCCATCAG | ACCACGCATA | CCGGCAAGCT | 8160 |
| GTGAGAAGTT | TGAGATGTTA | CCACGGGCTC | CAGAGTCCAT | CATCATAACG | ATTGGGTCTT | 8220 |
| TAGGATCTTG | GTTAGCAATC | AAGCGTTTCT | CAAGTTTTTC | ACGGGCAGCA | CGCCATTGAG | 8280 |
| CTGTAACAGC | ATTGTAACGC | TCGTCGTCTG | TGATCATACC | ACGACGGAAT | TGTTTGGTGA | 8340 |
| TTTGTTCGAC | ACGTTTGTGT | GATTCTTCAA | TGATTTGAGC | CTTGTATCA | ACGACTGGGA | 8400 |
| TATCGGCAAT | ACCCACTGTC | AATCCTGCAA | GAGTTGAGTG | GTGGTAACCG | AGGTTCTTCA | 8460 |
| TGCGGTCAAG | TAGGGCAGAA | GTTTCTGTG | TACGGAAACG | TTTGAAGATT | TCAGCGATGA | 8520 |
| TATTTCCAAG | GTTTCTCTTC | TTGAATGGAG | GGTTGAGCTC | AAGATTGCTG | ATAGCTTCTT | 8580 |
| TGATATCTCC | ACCAAGTGGC | AAGAAGTATT | TAGCTGGAAC | ACCTTCTGTC | AAGTTGGCAT | 8640 |

811

| | | | | | | |
|------------|------------|-------------|------------|------------|-------------|-------|
| TGTTTGGTTC | TTGCAAGTAT | GGTAGCCCT | CTGGCATGAT | ATCGTTGAAG | AGAATTTTAC | 8700 |
| CAACTGTTGT | AAGCAAGACC | TTATGTCTTT | GCTCTTCTGT | CCAAGGCTTG | TTGAGGCTGT | 8760 |
| CTGTTGCGAT | ACCAACACGT | GAGTGGAGGT | GAACATAACC | ATTGCGGTAA | GCCATAACCG | 8820 |
| CTTCGTCACG | GTCTTTGAAG | ACCATTCCCTT | CACCTTCGCG | ACCAGCTTCT | TCCATGGTCA | 8880 |
| AGTAGTAGTT | ACCCAAAACC | ATGTCCTGAG | ATGGAGTAAC | TACCGGTTTC | CCATCTTTCTG | 8940 |
| GGTTCAAGAT | GTGCTCAGCA | GCTAGCATGA | GGATACGAGC | TTCTGCTTGT | GCTTCTTCTG | 9000 |
| AAAGTGGTAC | GTGGATGGCC | ATTGGTCCC | CGTCAAAGTC | AGCATTGTAG | GCTTCACAGA | 9060 |
| CAAGTGGGTG | CAAGCGAAGA | GCCTTACCAT | CAATCAAGAC | TGGCTCGAAG | GCTTGGATAC | 9120 |
| CCAAACGGTG | AAGGGTCGGT | GCGCGGTTCA | AAAGCACTGG | GTGTTCTTTA | ATCACTTCTT | 9180 |
| CAAGGATATC | CCAGATACGC | TCATCTCCGC | GTTCCACCAA | GCGTTTAGCT | GCTTTGACGT | 9240 |
| TTTGACGAT | ATCACGGGCA | ACGATTTCAC | GCATGACAAA | TGGTTTAAAG | AGTCAATCG | 9300 |
| CCATTTACAG | CGGCACACCA | CATTGGTACA | TCTTAAGAGT | TGGACCAACG | GCGATAACTG | 9360 |
| AACGTCTGA | GAAGTCAACA | CGTTTACCGA | GCAAGTTTGT | ACGGAAGCGT | CCTTGTTTAC | 9420 |
| CTTTAAGCAT | GTGGCTCAAT | GATTTCAATG | GACGGCTACC | TGGTCCTGTG | ATTGGACGAC | 9480 |
| CACGACGACC | ATTGTCAATC | AAAGCGTCAA | CTGCTTCTTG | AAGCATACGC | TTCTCATTTT | 9540 |
| GAACGATGAT | ACCTGGTGCA | TTTAACTCAA | GCAAACGAGC | CAAACGGTTG | TTACGGTTGA | 9600 |
| TAACACGGCG | GTAAGGTCA | TTCAAGTCAG | ATGAGGCAAA | ACGGCCACCA | TCCAAGTGCA | 9660 |
| ACATTGGACG | AAGATCTGGT | GGGATAACCG | GAAGGATGTT | AAGAATCATC | CATTCAAGTT | 9720 |
| TGTTTCCAGA | CTTGTAAGG | GCATCCAAAA | CATCCAAACG | ACGGATGGCT | TTGACACGCT | 9780 |
| TTTGTCAGT | AGCTGTTTTC | AATTCCTCTT | TGAGTTCAGC | AATTTCTTTT | TCAAGATCTA | 9840 |
| CTTGCTTCAA | AAGGTCTTGG | ATGGCTTCCG | CACCCATCTT | GGCAACAAAT | GAACCATAAC | 9900 |
| CATATTCACG | CAAGCGCTCT | CGGTATTTCG | GCTCTGTCAT | GATAGACTTG | TGCTCAAGTG | 9960 |
| GTGTATCCTT | AGGATCAATC | ACCACATAAG | CCGCAAAGTA | GATAACTTCC | TCGAGGGCAC | 10020 |
| GAGGGCTCAT | ATCAAGGTC | AAGCCCATAC | GGCTTGGAAT | CCCCTTGAAG | TACCAGATGT | 10080 |
| GAGATACAGG | AGCTTTCAAT | TCGATATGTC | CCATACGCTC | ACGACGAAC | TTCGTACGCG | 10140 |
| TTACTTCAAC | CCCACAGCGG | TCACAAACAA | TTCCTCTGTA | ACGAATGCGT | TTGTACTTAC | 10200 |
| CACAAGCACA | TTCCAGTCT | TTTGTAGGAC | CAAAGATCAC | TTCATCAAAG | AGTCCTTCAC | 10260 |
| GTTCTGGTTT | CAAGGTACGA | TAATTGATTG | TTTCAGGTTT | TTTGACTTCT | CCATAAGACC | 10320 |
| ATGAACGGAC | TTTACTTGGA | GAAGCTAGGG | TGATTTGCAT | ACTTTTAAAA | CGATTTACAT | 10380 |

812

| | |
|--|-------|
| CAACCACTAT TTCTTCCCTT TCTATTCTAA GTGAACCTGCT TATTCTTGTT CAGCAGCTTC | 10440 |
| TTCTGTGCT TCCGCTTTTG TTGCTTCTC AGCTTCTTCA GCTTCAAAGG CTGCTTTAGC | 10500 |
| CTCTGGGCT GCTTTTTCGC GGGCTTTTTC AAGGTCATCT ACGTGGATGA CATCTTCGTC | 10560 |
| CATTCCCTCA TCCAAGTCGC GAAGTCCAC TTCTTGGTCA TCTTCGTCTA GGACACGCAT | 10620 |
| GTCAAGACCA AGAGATTGCA ATTCTTTGAC AAGAATCTCG AAGGATTCTG GAACACCTGG | 10680 |
| TTTGTGAATT GGTTCGCTT TTGTAATAGC TTCATAGGCT TTCAAACGTC CGTTGATATC | 10740 |
| GTCCGACTTG TAAGTCAAGA TTTCTTGAAG GACATTTGAC GCACCGTAGG CTTCAAGAGC | 10800 |
| CCAAACCTCC ATCTCACCGA AACGTTGTCC ACCAACTGA GCCTTACCTC CGAGTGGTTG | 10860 |
| TTGGGTAACA GTTGAGTATG GTCCGACTGA ACGCGGTGC AATTTATCAT CAACCATGTG | 10920 |
| GTGGAGTTG ATCATGTACA TGAATCCGAC AGAAACACGG TTATCAAACG GTTCACCAGT | 10980 |
| ACGTCCATCG TAAAGGATCG TTTTGGCATC GCTATCCATA CCGCTTCTT TAACAGTTGA | 11040 |
| CCAAAGATCT TCAGAACTTG CTCCATCAA GACTGGTGA GCGATGTGAA TACCAAGAGT | 11100 |
| ACGAGCTGCC ATACCAAGGT GAAGCTCCAT AACCTGACCG ATATTCATAC GTGATGGTAC | 11160 |
| CCCAAGTGGG TTCAACATGA TGTCGACTGG AGTTCCTCT GGAAGGTAAG GCATGTCTTC | 11220 |
| TACAGGAACG ATACGAGAGA CAACCCCTTT GTTCCGCTGA CGTCCGGCCA TTTTATCTCC | 11280 |
| GACCTTAATC TTACGTTTTT GAGCGATGTA AACACGAACC AACATGTAA CACCTGATTG | 11340 |
| CAACTCATCT CCATTTACAC GTGTAAAGAT CTTAACATCA CGAACGACAC CATCGGCACC | 11400 |
| GTGTGGTACA CGAAGAGAAG TATCACGCAC TTCACGAGAC TTGTCTCAA AGATAGCGTG | 11460 |
| CAAGAGACGT TCTTCAGCTG AAAGATCTTT CTCACCTTA GGTGTTACTT TACCTACAAG | 11520 |
| AATATCACCT TCTTTAACCT CAGCACCAAT ACGGATAATC CCCATTTCTG CAAGGTCTTT | 11580 |
| GAGGGCATCT TCACCAACGT TTGGAATTC GCGAGTGATT TCTTCAGGCC CAAGCTTTGT | 11640 |
| ATCGCGCGTT TCTGATTCGT ATTCTTCAAG GTGAACAGAT GTGTAGACAT CGTCTTCAC | 11700 |
| CAAGCGTTCG CTCATGATAA CGGCATCTC GAAGTTGTAA CCTTCCAAG TCATGTAGGC | 11760 |
| AACGATTGGG TTTTGTCAA GCGCCATTT TCCATTTTCC ATAGAAGGTC CGTCAGCGAT | 11820 |
| GAAATCGCCT TTTTCAACGA CATCACCAAC TTTTACGAGA GTGCGTTGGT TGTAAGCAGT | 11880 |
| ACCTGAGTTT GAACGACGGA ATTTTGGAT GTGGTAAACA TCCAATGAAC CATCTTCACG | 11940 |
| ACGAATCTCT ACCTTGTCAG CATCTGCGTA AGTAACCTTA CCATCATACT GAGCAATCAC | 12000 |
| AGCCGCACCA GAATCGTGGG CTGCTTGGTA TTCCATACCA GTACCAACGT AAGGTGCCTG | 12060 |
| AGGATTAATC AATGGCACAG CCTGACGTTG CATATTGGCT CCCATGAGGG CACGGTTGGA | 12120 |
| GTCATCGTTT TCCAAGAAAG GAATACATGC TGTCGCAACG GCAACTACCT GTTTTGGTGA | 12180 |

| | |
|---|-------|
| AACGTCCATG TAGTCAACAA TATTAGCTGG ATACTCTTGG TTGACCCCTT GGTGACGTCC | 12240 |
| CATGACAATC TTCTCAGCAA AGGTTCATC TTCATTGAGA CGAGAGTTAG CCTGAGCTAC | 12300 |
| AGTATATTCA TCTTCTTCAT CAGCTGTCAA CCAAACAATT TCGTTCGTGA CAACACCTGT | 12360 |
| TTACCGGTCA ACCTTACGGT ATGGTGTGTT AACAAAACCA TATTTGTTCA AGTGTCCATA | 12420 |
| AGATGACAAG TTATTGATCA AACCGATGTT AGGTCCTTCA GGTGTCTCGA TTGGACACAT | 12480 |
| ACGACCATAG TGAGTGTAGT GCACGTCACG TACTTCATAT CCAGCACGGT CACGAGTCAA | 12540 |
| ACCACCAGGT CCTAAGGCTG ACAAACGGCG TTTGTGAGAC AACTCAGAAA GCGGGTTGTG | 12600 |
| TTGGTCCATG AACTGTGACA ACTGTGATGA ACCAAAGAAT TCTTTAACTG CAGCTGTTAC | 12660 |
| AGGACGGATA TTGATAATTT GTTGTGGTGT CAAGACTTCA TTGTCTGAA CAGACATACG | 12720 |
| TTACCGGACA TTACGTTCCA TACGAGAAAG TCCCAAACGT ACTTGGTTGG CAAGCAATTC | 12780 |
| ACCAACCGCA CGGATACGAC GATTTCCAAG GTGGTCGATA TCATCTACAC GGCCAAGTCC | 12840 |
| TTCAGCCAAG TTGAGGAAGT AGCTCATCTC AGCAAGGATA TCTGCAGGAG TCACCGTACG | 12900 |
| AACCTTGTC A TCTGGGTTAG CATTACCAAT GATCGTTACG ACGCGATCTG GATCAGTTGG | 12960 |
| AGCAATAACC TTGAATTTTT GAAGAACAAC AGGCTCAGTC ACAACGGCTG CATCGTTTGG | 13020 |
| GATGTAGACA ATCTTGTTCA AGTCGCCATC CAAATGGCTT TCAATGCTTT CAATCACGCT | 13080 |
| ACGAGTCATA ATCGTACCAG CTTCTACCAA GATTTCTCCA GTTTCAGGGT CTACCAATGG | 13140 |
| CTCTGCAATG GTTTGGTTGA GCAAACGTGT TTTAACATTG AGTTTTTTAT TGATTTTGTA | 13200 |
| ACGACCAACT GCTGCCAAGT CATAACGACG TGGGTCAAAG AAGCGAGCTA CAAGCAAGCT | 13260 |
| ACGTGAGCTT TCAGCCGTCT TAGGCTCACC TGGACGAAGG CGTTCGTAAA TTTCTTTCAA | 13320 |
| GGCTTCGTCT GTACGAGAGT CCATTGGATT CTTGTGGATA TCTTTTTCAA CAGTGTGCG | 13380 |
| AACCAATTCTG CTGTCACCAA AGATATCAAA GATTTTCATCA TCACCTGAGA AACCAAGAGC | 13440 |
| ACGAACCAAG GTTGTAATG GAATCTTACG AGTACGGTCG ATACGAGTGT AGGTGATATC | 13500 |
| TTTTGAGTCG CTTTCAAGTT CCAACCAAGC TCCACGGTTA GGGATAACAG TTGAACCATA | 13560 |
| GCCCACCTTA CCATTTTTGT CTACTTTGTC GTTAAAGTAA ACACCTGGTG AGCGGACCAA | 13620 |
| CTGAGAAACG ATAATACGTT CACCACCATT GATGATGAAA GTACCCATTT CTGTCATGAT | 13680 |
| TGGGAAATCA CCAAAGAAAA CTTCTTGGGT CTTGATTTCTG CTTGTTTCTT TATTGATCAA | 13740 |
| ACGGAAGGTT AAAAAATTG GTGCTGAGTA GCTAGCATCG TGGATACGAG CTTCTTCTAG | 13800 |
| CGTATATTTT GGTTCTTGA TTTCATATCC AACAAATCC AACTCCATTG TGTCTGTGAA | 13860 |
| GTTTGAAATT GGCAATACAT CTTCAAACAC TTCCTTAAGA CCGTGGTCTA GGAAAGCTTT | 13920 |

814

| | |
|--|-------|
| GAATGAGTCA GTTTGAATTT CAATCAAATT TGGTAAGTCA AGAACTTCTT TGATTCTTGA | 13980 |
| AAAAC TACGA CGGGTACGAT GTTTCCCGTA TTGAACGTCA TGTCCTGCCA AGATGATTCT | 14040 |
| CCTTTGTAAA TAAGTTCCAA GCCTTGTCAA TCAGGCTTTT CTAATCGTCA TATGGTTGTA | 14100 |
| AACCCCTTAT CACCGTGTCC TCTTGACGAA TTTTCAGAAT CTTAAGCCT CTGTTACAAA | 14160 |
| TGCTCAAAAT CTTGAAAAA AGCACAAAA GAGCAGCTAA ATCTGACTTT TTCAGAAGAT | 14220 |
| TTAACTGCTG TGAGCCTTGT CTGGACAATA TTTCAGACAA AACCTACGAC AAATGATTAC | 14280 |
| CCATATTATA CCCTATTTAG CTAGATTTTT CAAGGGGTTT CAGTAGGTTT TTGTTAAATT | 14340 |
| TTTTCCCAT GAAACTTGG CATCACATTC GAATCACGCT ATGGTACAAA AACTGAAAA | 14400 |
| AACTATTGAC TGAAATCAT TTTCAAGGTA TAATAATAAA CGTTAAGGCG GTATAGCCAA | 14460 |
| GTGGTAAGGC ACGGCTCTGC AAAAGCTTGA TCGTCGGTTC AAATCCGTCT ACCGCCTTCT | 14520 |
| ATAACTPGAT TTATCAGGT TCAAATGAAC AGAAAGCCCA ATTTGAAGGG CTTTTTTTAT | 14580 |
| TTTCCCTCGA ATAAATACGT ATAACTTAA AAACCTTTGG AGCGAGTTTG TGGCAGAGTT | 14640 |
| CTTCCATGG CATAATCCC TTTTGAAATC AG | 14672 |

(2) INFORMATION FOR SEQ ID NO: 112:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7902 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 112:

| | |
|---|-----|
| AGGAGACTAT TCAAGCCCAA ATTgAGTAGC CCAGCAAAGA CTGTATAGAC TGTGATACGT | 60 |
| TTTTTCATAGC CATTTGGTAAA GAGAATTTGG GAACCAAGAA TGGTATCTAA GGCCAGGATA | 120 |
| ATCGTACGAA AAGCGAAGAG AGAGGTCAAG ATGCCGCCCTC CGATATATTT TTCACTACCG | 180 |
| TAAAGTAGGA TGGCATTTTGG TCCTAAAACC ATGAGTCCAA AACTCAGTGG AATGATAAAG | 240 |
| AAGTTAAAGA TTCGACTACC TCTATTAACC AGAGAAACAT AGGCTTCTTT GTCTCCTTTC | 300 |
| CCCAGATAGT AACTGAGACG AGGCACACTC ACTCCAATTG CACCTGTTAC AACCCCAGCT | 360 |
| ATAACGGTCA CAATTCGCTG AGCTATGGTA TAGTAACTAA CGTTGACATC AATCCCTGTT | 420 |
| TTAACGAGGA AGAGGCGATC TAAAAAAGTG AAGAGCATAT TGGCATTGGC AAAGACTAAC | 480 |
| ATGGCTGTCA GAGGGAGAAA GAGTGGTTTA AAATCACTTA GGTGAATTTT AACAAGTTTG | 540 |
| ATGTCTCTTT TAATCCAAAA ATAAC TAATC AGGTAGTTAA TCAGCGTCGA TAACTCATC | 600 |
| ACAAGTGTAT AGACAACAAT ATCGTGTTC TTTTAAACA ATAAGAAAAT AGAGACCAGC | 660 |

815

| | |
|--|------|
| ATCAGGATAC GGATGAAGGC AGTTTGTGTA AAGAGAAAAC TGTAATTTTC CAGAGCTTCA | 720 |
| TTGACCCATT CGATTGAAAA AATCTGGGCA ATGAGTTGAA TCCCCATAAC AAGGTAGACC | 780 |
| TTTTTGTGCGA TTGGATTATC AGTAAAGAAG AGAGGATAGG CTAGGATATA GACAGCAGTG | 840 |
| GTCAAAATCG TACAAGCGAT GCACAAATAA AAAAGACTAG AAAAGGTTCT GTTAAGATCT | 900 |
| TTTTTGTAT CCTTGACATT ACTGATAGCC CTTAAACCGT AGTTATAGAC ACCATAAGTT | 960 |
| GCAAAGGGCA AGAAAAATGA CAAATAGTG TCGACTGAGT TGAAGTAACC ATAGTCAGTT | 1020 |
| CGGTCCAAGA CACGCGCGAC ATAGGTTCCA GTTAGGATGG GAAAAATAAT ATTCAAGACA | 1080 |
| CGAATTTCCA TGTAAGATAG AGCATTAAAT TTTATACTTT TCATTCAATT TACCTCGTTT | 1140 |
| TTTATTATAT CATAAAGTTA GCTAATAAGA AATGAAGGGC AGTAAGTCAA GTAATCACTT | 1200 |
| TGAAGTTTCA AATCTTAAGT TTTAAGTTT CTTAAGGAA AGTATATTAT TCTGAAGGAC | 1260 |
| TCTAAAATTT CGCAGCCATT TATTAGTAAT TGCTACAGAA TTCCTAGTCA TTACTAGAAA | 1320 |
| TGGACTAGTT TCTTTGAATA ATAGAACTGC ATAATTCTCC TATTCTAGAA GGGGAGGACC | 1380 |
| AGTATTTCTT TTATGATAGG ACTAGATTGT GGTATAATAG AGAGAATAAG TTTTTTTAGT | 1440 |
| AAGACAAAGG AGAAAATAGA TGATTTATGC AGGAATCTT GCCGGTGGAA CTGGCACACG | 1500 |
| CATGGGGATC AGTAACTGC CAAAACAATT TTTAGAGCTA GGTGATCGAC CTATTTTGAT | 1560 |
| TCATACAATT GAAAAATTG TCTTGAGCC AAGTATTGAA AAAATGTAG TTGGTGTCA | 1620 |
| TGGAGACTGG GTTCTCATG CAGAAGATCT TGTAAGTAAA TATCTTCCTC TTTATAAGGA | 1680 |
| ACGTATCATC ATTACAAAGG GTGGTGCTGA CCGCAATACA AGTATTAAGA ACATCATTGA | 1740 |
| AGCCATTGAT GCTTATCGTC CGCTTACTCC AGAGGATATC GTTGTACCC ACGATTCTGT | 1800 |
| TCGTCCATTT ATTACACTTC GCATGATTCA GGACAATATC CAACTTGCCC AAAATCATGA | 1860 |
| CGCAGTGGAC ACAGTGGTAG AAGCGTTGA TACTATCGTT GAAAGTACCA ATGGTCAATT | 1920 |
| TATTACAGAT ATTCCAAATC GTGCTCACCT TTATCAAGGA CAAACACCTC AAACATTCCG | 1980 |
| TTGCAAGGAC TTCATGGACC TTTATGGATC TCTTTCTGAT GAAGAGAAGG AAATCTTGAC | 2040 |
| AGATGCATGT AAAATCTTTG TGATCAAAGG AAAAGATGTG GCTTTGGCCA AAGGTGAATA | 2100 |
| CTCAAATCTG AAGATTACAA CCGTAACAGA TTTGAAGATT GCAAAAAGTA TGATTGAGAA | 2160 |
| AGACTAGTAA AATGATTAAT CAAATTATC AACTAACTAA GCCTAAGTTT ATCAATGTCA | 2220 |
| AATATCAGGA AGAGGCTATT GACCAAGAGA ATCATATCCT TATCCGTCCC AACTACATGG | 2280 |
| CTGTCTGTCA TCGGATCAG CGTTACTATC AGGGAAAACG TGATCCCAAG ATTTGAATA | 2340 |
| AAAAGCTTCC AATGGCAATG ATTCACGAGT CATGTGGAAC CGTCATTTCT GACCCGACCG | 2400 |

816

| | |
|--|------|
| GAACCTACGA GGTTCGGTCAA AAAGTTGTCA TGATTCCCAA TCAGTCTCCT ATGCAGAGTG | 2460 |
| ATGAAGAATT CTATGAAAAC TACATGACAG GGACCCATTT CTTGTCTAGT GGATTTGATG | 2520 |
| GCTTTATGAG AGAGTTTGT TCTCTCCCTA AAGATCGTGT GGTGGCTTAT GATGCTATTG | 2580 |
| AAGATACGGT TGCAGCCATT ACAGAGTTTG TCAGTGTGGG CATGCACGCT ATGAATCGTC | 2640 |
| TATTGACTCT TGCTCATAGC AAGCGGGAGC GGATCGCCGT TATTGGAGAT GGAAGTTTAG | 2700 |
| CTTTTGTGGT TGCCAATATT ATCAACTATA CTTTGCCAGA AGCAGAGATT GTGGTTATTG | 2760 |
| GTCGTCATTG GGAAAAGTTG GAACTCTTCT CATTTGCCAA AGAATGCTAT ATTACGGATA | 2820 |
| ATATTCCTGA AGATTTGGCC TTTGACCATG CTTTGAATG TTGTGGTGGT GATGGTACTG | 2880 |
| GACCAGCTAT TAATGACTTG ATTCGCTACA TTCGTCCTCA GGAACGATT CTCATGATGG | 2940 |
| GAGTTAGCGA ATATAAAGTC AATCTCAATA CTCGCGATGC CTTAGAAAAG GGCTTGATTT | 3000 |
| TGGTTGGGTC ATCTCGTTCT GGTGCGATTG ATTTTGAAAA TGCTATCCAA ATGATGGAAG | 3060 |
| TCAAGAAATT TGCCAATCGT CTTAAAAATA TCCTTTATCT AGAAGAACCT GTAAGAGAAA | 3120 |
| TTAAAGATAT TCATCGTGTC TTTGCAACCG ATTTAAACAC AGCCTTTAAA ACAGTGTTTA | 3180 |
| AGTGGGAAGT ATAAGTACTG GAGGTTAATT GTGAGAAAA TCATTAAAGA AAAAATTCT | 3240 |
| TCCTTACTTA GTCAAGAAGA GGAAGTCCTC AGTGTGAAC AACTGGGTGG AATGACCAAT | 3300 |
| CAAACTATT TGGCCAAAAC AACAAATAAG CAATACATTG TTAAATTCTT TGGTAAAGGG | 3360 |
| ACAGAAAAGC TTATCAATCG ACAAGATGAA AAGTACAATC TTGAACTACT AAAGGATTTA | 3420 |
| GGCTTAGATG TAAAAAATTA TCTTTTGAT ATTGAAGCTG GTATCAAAGT AAATGAGTAT | 3480 |
| ATCGAATCTG CGATTACGCT TGATTCAACG TCAATCAAGA CCAAGTTCGA CAAAATTACT | 3540 |
| CCAATATTAC AAATATTCA TACGTCTGCT AAGGAATTAA GAGGAGAATT TGCTCCTTTT | 3600 |
| GAAGAAATCA AAAAATACGA ATCCTTGATT GAAGAACAAA TTCCTTATGC CAACTATGAA | 3660 |
| TCTGTTAGAA ATGCAGTCTT CTCCTTAGAG AAAAGACTGG CTGACTTAGG TGTGACAGA | 3720 |
| AAATCTTGTC ATATCGATTT GGTGCCTGAA AACTTTATCG AATCACCTCA AGGACGACTT | 3780 |
| TATTTGATG ACTGGAATA TTCATCAATG AATGATCCAA TGTGGGATTT GGCTGCCCTC | 3840 |
| TTTTTAGAGT CTGAATTCAC TTCCAAGAG GAAGAACTT TCTTATCTCA CTATGAGAGT | 3900 |
| GACCAAACAC CGGTTTCTCA TGAAAAGATT GCTATTTATA AAATTTTACA AGATACTATT | 3960 |
| TGGAGTCTAT GGAAGTCTA TAAGGAAGAG CAAGGTGAAG ATTTGGTGA CTATGGTGTG | 4020 |
| AATCGTTACC AAAGAGCTAT TAAAGGTTTG GCTTCTTATG GAGGTTTACA TGAAAAGTAA | 4080 |
| AAACGGAGTT CCTTTGGCC TTCTCTCAGG TATTTTCTGG GGCTTGGGTC TAACGGTTAG | 4140 |
| TGCTTATATC TTTTCGATTT TTACAGATTT GTCACCCTTT GTGGTGGCTG CAACTCATGA | 4200 |

| | |
|---|------|
| TTTTTTGAGC ATCTTTATCT TACTAGCTTT TCTCTTGGTA AAAGAAGGGA AAGTTCGCCT | 4260 |
| CTCAATTTTC TTAAATATTC GCAATGTCAG TGTATCATC GGAGCCTTGC TAGCAGGCCC | 4320 |
| TATCGGTATG CAGGCCAATC TTTATGCAGT TAAGTATATC GGAAGTTCTT TAGCTTCATC | 4380 |
| TGTATCGGCT ATTTACCCTG CGATTTCAGT TCTATTGGCT TTCTTCTTTT TGAAGCACAA | 4440 |
| GATTTGAAA AATACTGTAT TTGGGATTGT CTTGATTATT GGAGGGATTA TTGCTCAGAC | 4500 |
| CTATAAGGTT GAACAGGTTA ATTCTTTCTA CATTGGGATT CTTTGTGCTT TGGTTTGTGC | 4560 |
| TATTGCATGG GGAAGTGAGA GTGTTCTTAG CTCTTTTGCC ATGGAAAGTG AATTGAGTGA | 4620 |
| AATCGAAGCC CTCTTAATCC GTCAAGTAAC TTCGTTCTTG TCCTATCTTG TGATTGTGCT | 4680 |
| CTTCTCTCAT CAGTCATTTA CTGCAGTAGC CAATGGACAA TTGCTAGGTC TCATGATTGT | 4740 |
| TTTTGCAGCC TTGATATGA TTTCTACTT GGCTTATTAT ATCGCTATCA ATCGCTTGCA | 4800 |
| ACCAGCCAAG GCTACAGGCT TGAACGTGAG CTATGTAGTA TGGACGGTCT TGTTCGAGT | 4860 |
| TGTTTTCTTG GGTGCACCGC TAGATATGCT GACCATTATG ACGTCACTTG TCGTCATTGC | 4920 |
| TGGAGTTTAT ATTATTATTA AAGAATAAAG GAGATTCGTG TGAAAGCCAT TATCTTAGCA | 4980 |
| GCGGGATTGG GAACTCGCTT GCGTCCTATG ACTGAAAATA CCCCTAAAGC CTTGGTTCAG | 5040 |
| GTTAATCAAA AACCTTTGAT TGAGTACCAA ATTGAGTTTC TCAAAGAAAA AGGAATCAAT | 5100 |
| GACATCATCA TCATTGTTGG TTATCTTAAA GAACAATTCG ATTACTTGAA AGAGAAATAC | 5160 |
| GGTGTTCGTC TCGTTTTCAA TGATAAATAC GCTGACTACA ATAACTTTTA CTCTCTCTAT | 5220 |
| CTTGTAAGAG AAGAATTGGC CAACAGCTAT GTTATTGATG CTGACAATTA TCTCTTAAA | 5280 |
| AATATGTTCC GCAATGATTT GACACGTTTC ACTTATTTTA GTGTTTATCG TGAAGATTGT | 5340 |
| ACCAACGAAT GGTTCCTGGT TTATGGAGAT GACTACAAGG TTCAAGACAT TATTGTTGAT | 5400 |
| AGCAAGGCAG GTCGCATCCT TAGTGGTGTA TCCTTCTGGG ATGCTCCAAC TGCAGAAAAG | 5460 |
| ATTGTCAGCT TTATCGACAA GGCTTATGTA AGTGGTGAAT TTGTTGATCT CTATTGGGAC | 5520 |
| AATATGGTTA AGGATAATAT CAAAGAGCTA GATGTCTATG TTGAAGAATT AGAAGGCAAT | 5580 |
| AGCATTATAG AGATCGATAG TGTCCAAGAC TATCGTAAAT TAGAAGAAAT TCTTAAAAAC | 5640 |
| GAAAATTAAA GATTCCAACA TCTGACAAA TAGTCGGATG TTTTTTGATT TTTTACGAAC | 5700 |
| TTTTACGAAT AGATAGATGA GTAGAAAAAG AAATGGAGTT ATTTATGAAA ATCACAACCT | 5760 |
| ATGAAATCTA TAAGTTAAAA AAATCAGGTT TGACCAATCA ACAGATTTTG AAAGTGCTAG | 5820 |
| AATACGGTGA AAATGTTGAT CAGGAGCTTT TGTGGGTGA TATTGCAGAT ATCTCAGGTT | 5880 |
| GCCGTAATCC AGCCGTTTTT ATGGAACGTT ATTTTCAGAT AGACGATGCG CATTTGTGCA | 5940 |

818

| | |
|--|------|
| AAGAGTTTCA AAAATTTCCA TCTTTCTCTA TTTTAGATGA CTGTTATCCT TGGGATTTGA | 6000 |
| GTGAAATATA TGATGCGCCT GTACTTTTAT TTTACAAGGG AAATCTTGAC CTCCTGAAAT | 6060 |
| TCCCGAAGGT AGCGGTCGTG GGCAGTCGTG CTTGTAGCAA ACAGGGAGCT AAGTCAGTTG | 6120 |
| AAAAAGTCAT TCAAGGCTTG GAAAAATGAAC TGGTTATTGT CAGTGGTCTG GCCAAGGGCA | 6180 |
| TTGACACAGC AGCTCATATG GCAGCTCTTC AGAATGGCGG AAAAACCATT GCAGTGATTG | 6240 |
| GAACAGGACT GGATGTGTTT TATCCTAAAG CCAATAAAGC CTTGCAAGAC TACATCGGCA | 6300 |
| ATGACCATCT GGTCTAAGT GAATATGGAC CTGGTGAACA ACCTCTGAAA TTTCATTTTC | 6360 |
| CTGCCCGTAA TCGCATCATT GCTGGACTTT GTCGTGGTGT GATTGTAGCA GAGGCTAAGA | 6420 |
| TGCGTTCAGG TAGTCTCATT ACGTGTGAGC GAGCAATGGA AGAAGGACGC GATGCTTTTG | 6480 |
| CTATTCTCGG TAGCATTTTA GATGGACTAT CAGACGGTTG CCATCATTTG ATTCAAGAAG | 6540 |
| GAGCAAAATT GGTCAACAGT GGGCAAGATG TTCTTGCGGA ATTTGAATTT TAAAAATGAC | 6600 |
| CTAAGCTAGA ATTCTAAGAA AAAATCAATT TTAAGAGAAA ATGAACCCAA CATTTCCATA | 6660 |
| ATAAAACGCA TATTAGCAAG TTTTAAACAC TTGATAATAT GCGTTTTTTC TAAGTGGATT | 6720 |
| AGTAGAGTAG AGGATTTTTC TCAATATAA CTCTTCGAAA ATCTCTTCAA ACTACGTCAG | 6780 |
| CTTCCATCTG CAACCTCAAA ACAGTATTTT GAGCGaCTtC GTCAGTCTTA TCTACACCT | 6840 |
| CAAAGCAGTG CTTTGAGCAA CCTGTGGCTA GCTTCCTAGT TTGCGCTTTG ATTTTCATTG | 6900 |
| AGTATAAGGG AAAGTATAGT GAATTGAAAT AAGATGTGAA CAACTCTATC AGGAAAGTCA | 6960 |
| AATTAATTTA TAGAAATATT TTAGCAGCCA AGGTGTACTG TTATAGATTC AATTACACTA | 7020 |
| TAATTTAGTG TAATTGAGAA AGGAGAAATG ATTGTGATTG ATGTTGGCTA GGTATGTTC | 7080 |
| AATGATTCCCT ACCGTCTCAA ATCTTGTCAG TAAGGAAAAA TAAATTCTTC AAAAGTAGAG | 7140 |
| ATTACAAGGC TTGTTTAAGA AAGAATTCAA AGACCTTGAC AAATAAAAAT AAAATGGTTA | 7200 |
| TTATAAAAAA TGGTCTGAAA TAGATGATGA TACTTTTCGA AAATCTCTTC AAATACGTCA | 7260 |
| GCTCAGCTTT GCCTTGCTGT GTTTTGAGCA AGCTACGGTT AGCTTCCGAG TTTGATTTTC | 7320 |
| ATTACTAGA AATGAACTG ATGAGAGATA TCAGTAGACA TTTGAGTCAG GATATTATGG | 7380 |
| AAAATGATAA AAAGAGCTCG TGAGATTGGC ATATCAGACT ACTAAAGTAT TGAGTTGTTC | 7440 |
| AGGATTTTAG CGACTAGTTA GCTGGGAAAG GAAGATATTT GTGACAAATA ATAACTGTAT | 7500 |
| TTCGTTGATA GAATTTAGAA ATAAAAATATA TGAAGAATTA GAACTTTCCA GAAGTGATTT | 7560 |
| AGCGATTTTA CTATGTGCCA TGCTTATCGC CTCTATCGGA TTAAATATGG ATTCGACTCC | 7620 |
| CGTGATTATT GGAGCCATGT TAATCTCTCC TTTGATGACA CCTATTCTGG GAGTGGGGCT | 7680 |
| CTCTCTAGCT ATATTGATT TTAAATTGTT AAGAAAATCT TTAAAAATAT TAGCTATTCA | 7740 |

819

| | |
|---|------|
| AATTCTTGCC AGTCTAATAG CTTCACACT TTATTTTTAT CTTTCTCCCA TTTCGTATGC | 7800 |
| TAGTTCGGAG ATTGTTGCTA GAACCTCTCC GACTATTTGG GATGTTCTCA TTGCTTTTGT | 7860 |
| AGGAGGGGATA GCAGGTATCA TTGGTGCTAG GAAAAAAGAG AC | 7902 |

(2) INFORMATION FOR SEQ ID NO: 113:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 18627 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 113:

| | |
|---|------|
| GAAGTTGAAA TGGCCAGCTG ATGAGCAATA TCGGTCATAG AAATCTTCTC AATCAACTTT | 60 |
| TGCGCAATTT TTTGGTTGAT AATACGAGGA ATTTGGTGAT TTTTCTTGAC GATAGAAGTT | 120 |
| TCAGCGACCA TCATTTTGA ACAGTGATAG CACTTGAAAC GACGCTTTCT AAGTAGAATT | 180 |
| CTAGTAGGCA TACCAGTTGT CTCAGGTAA GGAATCTTAG ACGGTTTTTG AAAGTCATAT | 240 |
| TTCTTCAATT GGTTCGCA CTCAGGGCAA GATGGGGCGT CGTAGTCCAG TTTGGCGATG | 300 |
| ATTTCTTGT GTGTATCTT ATTGATGATG TCTAAATCT GGATATTAGG GTCTTTAATG | 360 |
| TCTAGTAATT TTGTGATAA ATGTAATTGT TCCATATGAA TCTTTCTAAT GAGTTGTTTG | 420 |
| GTCGCTTTTC ATTATAGGTC ATATGGGACT TTTTCTTAC AATAAAATAG GCTCCATAAT | 480 |
| ATCTATAAGG GATTACCCA CTACAAATAT TATAGAGCCA AAAATCCTTT GTTACTATAA | 540 |
| CAAGGGATTT TTCTTTTGT TCTGCTCCTT TTTTGATATA ATAGTTCTAT GTTAAATCA | 600 |
| GAAAAACAAT CACGTTATCA AATGTTAAAT GAAGAATTGT CCTTCCTATT GGAAGGCGAA | 660 |
| ACCAATGTTT TGGCTAATCT TTCCAACGCC AGTGCTCTCA TAAATCACG TTTTCTAAT | 720 |
| ACCGTATTTG CAGGCTTTTA TTGTTTCGAT GGAAAGGAAT TGGTTTATAG CCCCTTCCAA | 780 |
| GGAGTGTTT CCTGCATCCG TATTGCACTA GGCAAGGGTG TTTGTGGTGA GGCAGCTCAC | 840 |
| TTTCAGGAAA CTGTTATTGT TGGAGATGTG ACGACCTATC TCAACTATAT TTCTTGTGAT | 900 |
| AGTCTAGCTA AAAGTGAAAT TGTGGTGCCG ATGATGAAGA ATGGTCAGTT ACTTGGAGTT | 960 |
| CTGGATCTGG ATTCTTCAGA GATTGAGGAT TACGATGCTA TGGATCGAGA TTATTGGAA | 1020 |
| CAATTGTGCG CTATTTTGCT TGAAAAGACA GCATGGGACT TTACGATGTT TGAGGAAAAA | 1080 |
| TCTTAATGTA TCAAGCACTT TATCGAAAAT ATAGAAGTCA AAACCTCTCC CAGTTAGTTG | 1140 |
| GTCAAGAAGT TGTGGCTAAG ACTCTTAAAC AAGCGGTGGA GCAAGAGAAA ATAAGTCACG | 1200 |

820

| | |
|--|------|
| CTTATCTTTT TTCTGGTCCT CGTGGAACGG GAAAAACCAG TGTGCTAAA ATCTTTGCCA | 1260 |
| AGGCTATGAA CTGTCCCAAT CAAGTGGGTG GCGAACCTTG CAATAACTGC TATATTTGTC | 1320 |
| AAGCAGTGAC GGACGGTAGT TTAGAAGATG TCATTGAAAT GGATGCAGCT TCTAATAATG | 1380 |
| GGGTAGATGA AATTCGCGAA ATTCGTGATA AATCTACCTA TGCGCCTAGC CTTGCTCGTT | 1440 |
| ATAAGGTTTA TATCATAGAT GAGGTTTACA TGCTGTCTAC AGGGGCTTTT AATGCCCTCC | 1500 |
| TAAAGACGCT GGAAGAACCA ACACAGAATG TAGTCTTTAT TTTGGCCACT ACTGAATTGC | 1560 |
| ACAAGATTCC TGCTACTATT CTATCCCGTG TGCAACGTTT TGAGTTTAAA TCAATTAAGA | 1620 |
| CACAGGATAT TAAGGAACAT ATTCACTATA TCTTAGAAAA AGAAAATATC AGTTCTGAAC | 1680 |
| CAGAGGCTGT GGAAATCATT GCCAGACGGG CGGAAGGTGG AATGCGGGAC GCCTTGTCTA | 1740 |
| TTTGGATCA AGCCCTGAGT TTGACACAGG GAAATGAGCT GACGACTGCT ATCTCTGAAG | 1800 |
| AAATTACTGG CACCATTAGC CTATCAGCCT TGGATGATTA TGTGGCGGCC TTGTCTCAAC | 1860 |
| AGGATGTTCC CAAAGCTTTG TCTTGCTTGA ATCTTCTTTT TGACAATGGT AAGAGCATGA | 1920 |
| CTCGTTTGTG GACCGATCTT TTGCACTATT TAAGAGACTT GTTAATTGTT CAAACAGGGG | 1980 |
| GAGCAAATAC TCATCATAGT TCAGTCTTTG TAGAAAAATT GGCACCTTCCT CAAAAAATC | 2040 |
| TGTTTGAAAT GATTCGCTTA GCAACAGTGA GTTTAGCAGA TATTAAGTCT AGTTTGCAAC | 2100 |
| CCAAGATTTA TGCTGAAATG ATGACCGTCC GTTTGGCGGA AATCAAGTCC GAACCAGCTC | 2160 |
| TATCAGGAGC GGTGAAAAAT GAAATTGCTA CGCTGAGACA GGAAGTTGCC CGTCTCAAAC | 2220 |
| AAGAGCTTTC TAATGTAGGT GCGGTTCCCTA AACAAGTTGC ACCAGCTCCT AGTCGACCAG | 2280 |
| CTACGGGCAA AACAGTCTAT CGTGTGATC GCAATAAAGT GCAATCTATC TTACAAGAGG | 2340 |
| CCGTCGAAAA TCCTGATTTA GCACGTCAAA ATTTAATTCG TTTGCAGAAT GCCTGGGGAG | 2400 |
| AGGTAATTGA AAGTCTAGGT GGGCCGGACA AGGCTCTGCT AGTTGGTTCT CAACCGGTTG | 2460 |
| CTGCCAATGA ACACCATGCT ATTCTTGCTT TTGAGTCTAA CTTCAATGCT GGTCAAACTA | 2520 |
| TGAAACGAGA CAATCTCAAT ACCATGTTTG GTAATATCCT CAGTCAGGCG GCAGGTTTTT | 2580 |
| CACCTGAGAT TTTAGCTATT TCCATGGAGG AATGGAAAGA AGTTGCGGCA GCCTTTTCAG | 2640 |
| CCAAAGCCAA ATCTTCTCAA ACTGAAAAAG AAGTAGAAGA AAGCCTGATT CCAGAAGGAT | 2700 |
| TTGAATTTTT GGCTGATAAA GTGAAGGTAG AGGAAGACTA AAGAAAGATT TCATGATACA | 2760 |
| ATAAGTTTAT GAATAACAA CAATTTATTA TTATGGCGCT GTTTACAGCT GCTGAGACCT | 2820 |
| ATTTTTTCAA TGAAGCCTGG ATGACTGGCC GCTATATTAT GGCAGCCTTT TGGGCAATTT | 2880 |
| TACTCTTTAG AAATTTCCGA GTCAGTTATG TGATGGGCAA AATCGTTGAT GTCATCGATC | 2940 |
| AGCATTTTAA TAGGAAAGAC TAGCCCTCAG CTTCAGACA AAATCAAAGC CTTTTAGGCT | 3000 |

821

| | |
|--|------|
| TTTTTTTGT ATACTAGAAA AGTATATTTA TAGAATTTTT GCTCTATTTT TGGGGAAATC | 3060 |
| AGACGTTTTT CTAGTAAGTA CTGTAAAAGT TTTGAAAAAG AAAGGAAC TAATGTCAGT | 3120 |
| ATTAGAGATC AAAGATCTTC ACGTTGAGAT TGAAGGAAAA GAAATTTTAA AAGGGGTAA | 3180 |
| CCTGACCCTG AAAACAGGAG AAATTGCCGC TATCATGGGA CCAAATGGTA CAGGTAAATC | 3240 |
| GACTCTTTCT GCCGCTATCA TGGGAAATCC AAACATATGAA GTAACATAAG GTGAAGTTTT | 3300 |
| GT'TGATGGC GTAAACATCC TTGAGTTGGA AGTGGATGAG CGTGCGCGTA TGGGACTTTT | 3360 |
| CCTTGCTATG CAATACCCAT CAGAAATCCC TGGAATTACC AATGCTGAGT TTCTTCGTGC | 3420 |
| CGCTATGAAT GCGGGTAAAG AAGATGATGA GAAGATTTCA GTTCGTGAGT TTATTACTAA | 3480 |
| GCTAGATGAA AAAATGGAAT TGCTCAACAT GAAAGAAGAA ATGGCAGAGC GTTACCTCAA | 3540 |
| CGAAGGCTTC TCTGGTGGTG AGAAAAACG CAATGAAAT CTTCACCTTT TGATGTTGGA | 3600 |
| GCCAACATTT GCTCTTTTGG ACGAGATTGA CTCAGGCTTT GATATTGACG CTCTTAAAGT | 3660 |
| TGTGTCTAAA GGTGTCAATG CCATGCGTGG TGAAGGTTTT GGTGCTATGA TCATCACTCA | 3720 |
| CTACCAACGT CTTTGAAGT ATATCACACC TGATGTGGTA CACGTGATGA TGAAGGTCG | 3780 |
| TGTTGTCCTT TCTGGTGGTC CAGAATTGGC TGCGCGTTTG GAACGTGAAG GATACGCAAA | 3840 |
| ATTAGCTGAA GAACTGGCT ACGACTACAA GGAAGAATTG TAATCCCTC GTATCTTTTA | 3900 |
| GGAGAAGTAA ATGACTAGAG AAAATATTAA ACTTTTTTCA GAAATGCACG CTGAACCAAG | 3960 |
| CTGGTTGGCT GATCTCCGTC AAAAAGCTTT TGACAAGATT GAGACTTTGG AATTACCAGT | 4020 |
| TATTGAGTGT GTCAAATCC ACCGTTGGAA TCTGGGTGAT GGAACGATTA CAGAAAATGA | 4080 |
| GCCATCAGCA AATGTTCCAG ATTTACACAG TTTAGATCAT CACTTGAAGT TGGTGCAAGT | 4140 |
| AGGAACTCAA ACTGTTTTCG AACAACTCC AGTTGAGTTA GCTGAACAGG GTGTTGTCTT | 4200 |
| CACAGACTTT CACTCAGCTT TAGAAGAAAT TCCAGAGCTG ATCGAAGAAT TCTTCATGTC | 4260 |
| ATCTGTTAAG TATGATGATG ACAAGTTGGC GGCTTACCAC ACAGCTTACT TTAACAGTGG | 4320 |
| TGCTGTACTC TATATTCCAG ATAACGTAGA AATCACAGAG CCAATTGAAG GAATTTTCTA | 4380 |
| CCAAGATAGC GATAGCAATG TGCCGTTTAA CAAGCATATT ATGATTATCG TTGGTAAAAA | 4440 |
| TTCTAAGATT AGTTATCTGG AGCGTTTAGA GTCACGCGGT GAAGGAAGTG ACAAAGCAAC | 4500 |
| TGCCAATATC ACAGTGGAAG TGATTGCACG TTCTGGTGCG CAAGTCAAGT TTGCTGCTAT | 4560 |
| CGACCGTCTA GGTGAAAACG TCACTGCCTA CATTAGCCGT CGTGGTAAAT TAGGCAACGA | 4620 |
| TGCAAGTATT GACTGGGCTA TCGGTGTCAT GAACGAAGGA AATGTCGTTG CTGATTTTGA | 4680 |
| TAGTGACTTG ATGGTAATG GTAGCCATGC TGACCTCAAG GTTGTAGCTC TTTCAAGTGG | 4740 |

822

| | | | | | | |
|-------------|-------------|------------|------------|------------|------------|------|
| TCGTCAGGTA | CAAGGGATTG | ATACTCGTGT | AACTAACTAT | GGCTGCAACT | CAATCGGAAA | 4800 |
| CATTCTACAA | CATGGGGTTA | TCCTTGAAAA | AGCAACTTTG | ACTTTCAATG | GTATCGGCCA | 4860 |
| CATCATCAAG | GGTGCTAAGG | GAGCAGATGC | GCAACAAGAG | AGCCGTGTTC | TCATGCTTTC | 4920 |
| AGACCAAGCG | CGTTCAGATG | CTAACCCAAT | TCTTTTGATT | GATGAAAATG | ACGTAAGTGC | 4980 |
| AGGCCATGCA | GCCTCTATTG | GTCAGGTAGA | TCCAGAAGAT | ATGTACTACC | TCATGAGTCG | 5040 |
| TGGCTTGAT | AAGGCAACTG | CAGAGCGTTT | GGTTGTTTCG | GGTTTCCTTG | GATCTGTTAT | 5100 |
| CGTGGAGATT | CCAGTCAAGG | AAGTTCGTGA | TGAAATGATT | GCAACTATCG | AAGAGAAATT | 5160 |
| GTCAAAACGC | TAAGGGGCAG | CCTATGTTAG | ATGTAGAAGC | GATTCGCAAG | GATTTTCCAA | 5220 |
| TTTTAGATCA | GATTGTCAAT | GATGAACCTC | TGGTCTATCT | GGACAATGCT | GCGACGACAC | 5280 |
| AAAAACCACT | AGTAGTTCCTG | AAAGCTATTA | ACAGCTACTA | TGAGCAGGAC | AATGCCAATG | 5340 |
| TTCACCGTGG | TGTCCATACC | TTAGCGGAAC | GAGCGACAGC | TTCTTATGAA | GCTGCTCGTG | 5400 |
| AAACCATTCG | TAAGTTTATT | AATGCAGGCT | CTACAAAGGA | AGTTCTCTTT | ACCAGAGGAA | 5460 |
| CGACAACCAG | CCTTAACTGG | GTGGCACGCT | TTGCTGAGGA | AATTCTCACT | GAGGGAGACC | 5520 |
| AGGTCTTGAT | TTCAGTAATG | GAACACCATT | CTAATATCAT | TCCATGGCAG | GAAGCTTGTC | 5580 |
| GAAAGACTGG | AGCAGAGCTT | GTCTATGTCT | ATCTTAAAGA | CGGTGCCTTG | GATATGGAGG | 5640 |
| ATTTGCGAGC | TAAATTGACT | GATAAGGTTA | AATTTGTTTC | CCTAGCTCAT | GCCTCCAATG | 5700 |
| TTCTTGGTGT | GGTCAATCCG | ATCAAGGAAA | TCACTCAATT | AGCCCACCAA | GTTGGGGCAA | 5760 |
| TTATGGTAGT | GGATGGTGCT | CAATCTACAC | CTCATATGAA | GATTGATGTC | CAGGACTTGG | 5820 |
| ATCTGGACTT | TTTCGCCTTT | TCGGGTCACA | AGATGGCTGG | TCCGACTGGT | ATCGGTGTCC | 5880 |
| TTTACGGCAA | AGAAAAGTAT | CTTGAGCAAA | TGTCTCCAGT | AGAATTGCGC | GGCGAGATGA | 5940 |
| TTGATTTTGT | CTACGAGCAA | TTTGCTAGTT | GGAAGGAATT | GCCTTGGAAG | TTTGAGGCTG | 6000 |
| GAAACGCCAAA | TATGGCAGGA | GCTATTGGAC | TTGCGACTGC | AGTTGATTAT | CTGGAAAAGA | 6060 |
| TTGGTATGGA | TGCCGTTGAA | GCTCATGAAC | AGGAATTGAT | TGCGTACGTC | TATCCAAAAC | 6120 |
| TGCAGGCAAT | TGAGGGATTG | ACCATTTACG | GTTCTCAGGA | TTTGGCTCAA | CGTTCGGGTG | 6180 |
| TTATTGCCTT | TAACCTAGGT | GATCTCCATC | CTCACGATCT | TGCGACGGCT | CTGGATTATG | 6240 |
| AAGGAGTGGC | TGTTCTGTCT | GGTCACCATT | GTGCGCAACC | CTTGCTTCAG | TATTTGGAAG | 6300 |
| TCCCAGCAAC | AGCTCGTGCA | AGTTTTTATA | TCTACAATAC | CAAGGCAGAT | TGCGACAAAC | 6360 |
| TAGTCGATGC | CCTACAAAAG | ACAAAGGAGT | TTTTCAATGG | CACCTTCTAA | ACTAGATAGC | 6420 |
| CTTTATATGG | CAGTGGTAGC | AGACCATTGC | AAAAATCCAC | ATCACCAAGG | GAAGTTAGAA | 6480 |
| GATGCTGAGC | AAATCAGTCT | CAACAATCCG | ACTTGTGGGG | ATGTCATCAA | CCTCTCTGTC | 6540 |

823

| | |
|---|------|
| AAGTTTGATG CAGAGGACCG TTTGGAAGAT ATTGCTTTTC TAAATTCAGG ATGCACGATT | 6600 |
| TCAACTGCCTT CTGCTAGTAT GATGACAGAT GCCGTTT TAG GAAAAACCAA ACAAGAAATT | 6660 |
| TTAGAACTGG CGACTATTTT TTCTGAAATG GTTCAAGGGC AAAAAGATGA GCGTCAAGAC | 6720 |
| CAACTTGGAG ACGCGGCATT CTTGTCAGGT GTTGCCAAAT TCCCTCAAAG AATCAAGTGT | 6780 |
| GCAACCCTAG CTTGGAATGC CCTTAAGAAA ACAATTGAAA ATCAAGAAAA ACAGTAAGAC | 6840 |
| AAGTTTCTTT TGTCTTATGA ATTATTAGAA ATGAAGAAAG AAAGGATACT ATGGCTGAAG | 6900 |
| AAAGAGTAGA ACCAAAACCA ATTGACCTTG GTGAATATAA ATTTGGTTTC CATGACGATG | 6960 |
| TAGAGCCTGT CTTATCGACA GGAAAAGGAC TCAACGAAGG TGTATTTCGT GAATTATCTG | 7020 |
| CTGCTAAGGG TGAGCCTGAG TGGATGTTGG AGTTCCGTTT GAAGTCTTAT GAAACCTTCA | 7080 |
| AAAAATGCC CATGCAAACT TGGGGAGCAG ACTTGTCAGA GATTGACTTT GATGACTTAA | 7140 |
| TCTACTACCA AAAACCATCT GACAAACCAG CCCGTTCTTG GGATGATGTA CCTGAAAAGA | 7200 |
| TTAAAGAAAC CTTTGAACGT ATCGGGATTTC CAGAAGCTGA ACGTGCTTAT TTAGCAGGGG | 7260 |
| CTTCGCCCCA GTACGAGTCA GAAGTGGTTT ACCACAACAT GAAGGAAGAG TTCCAAAAAT | 7320 |
| TAGGTATTAT CTTTACAGAT ACAGATTCCG CACTCAAGGA ATACCCAGAC TTATTTAAAC | 7380 |
| AATACTTTGC GAAGTTGGTA CCGCCGACAG ATAACAAGTT GGCAGCCCTC AACTCAGCAG | 7440 |
| TATGGTCGGG TGGAACTTT ATCTACGTGC CAAAAGGTGT CAAGGTAGAT ATTCCACTTC | 7500 |
| AAACTTATTT CCGTATCAAT AACGAAAATA TAGGTGAGTT CGAACGTACC TTGATTATCG | 7560 |
| TTGATGAGGG AGCAAGCGTC TACTACGTAG AAGGATGTAC AGCACCAACA TATTCAAGCA | 7620 |
| ATAGCTTACA CGCTGCCATT GTAGAAATTT TTGCTTTGGA CGGAGCTTAT ATGCGTTATA | 7680 |
| CAACTATCCA AAACCTGGTCT GATAACGTCT ATAACCTGGT AACAAAGCGT GCTAAGGCTC | 7740 |
| AAAAGGATGC CACTGTTGAG TGGATTGATG GAACTTGGG TGCCAAAACG ACTATGAAAT | 7800 |
| ATCCATCTGT TTACCTTGAT GGAGAAGGAG CGCGTGGTAC CATGCTCTCT ATCGCCTTTG | 7860 |
| CTAATGCAGG GCAACACCAA GACACGGGTG CTAAGATGAT TCACAATGCT CCACATACCA | 7920 |
| GCTCGTCTAT TGTGTCTAAA TCCATCGCTA AAGGTGGAGG AAAGGTTGAC TACCGTGGAC | 7980 |
| AAGTCACCTT TAACAAGAAC TCTAAGAAAT CTGTTTCCCA CATTGAATGT GATACCATTA | 8040 |
| TCATGGATGA CTTGTCAGCA TCAGATACTA TTCCATTAA TGAAATTAC AACTCGCAAG | 8100 |
| TGGCTTTGGA ACACGAAGCC AAAGTATCTA AGATTTCAGA AGAGCAATTG TATTATCTCA | 8160 |
| TGAGCCGTGG ATTGTCAGAA TCTGAGGCAA CTGAAATGAT TGTCATGGGA TTTGTAGAAC | 8220 |
| CCTTTACAAA AGAACTTCCA ATGGAATACG CAGTTGAGCT GAACCGCTTG ATTAGCTATG | 8280 |

824

| | |
|---|-------|
| AAATGGAGGG ATCAGTTGGA TAAATTTGA TTTTATACTC TTCGAAAATC TCTTCAAACC | 8340 |
| ACGTCAGCAT CGCCTTACCG TATGTATGGT TwCTGAtTCG TCAGTTTCAT CTACAACCTC | 8400 |
| AAAACAGTGT TTTGAGCAAC tGCGGCTAGC TTCCTAGTTT GTTCTTTGAT TTTGAGTATT | 8460 |
| AGATTTACTC AAAATCAAGG ATTTTGAAGA TGAACCTGTA TCAAAAAATC GCGGTTTAAA | 8520 |
| ATCGCGATTT TTTATAATTT CTCGTTAACA AAGCGGACAA ACTGATTCCA CCAAACTTT | 8580 |
| AAGAAGAAGG CTTTTTCAAT TTTCTTGTCT GCTACCATTT CGAAACTAGG GCGCTCTGTG | 8640 |
| GTGATGTAAC CTTGACCAAT CAAGTCCTTG TCTTCATAAG TCAAATGGCC AACCACGT | 8700 |
| CCAGCTTCAA GTGGTGCTGG GATTGCTTTG GAATCAGGTG TGAATTGAAC AGATTGGGAA | 8760 |
| GATTGATTCC CAACACGTC GATTAGATAG ATATCCTCTG GAGCCACTGC AGTTACTGTA | 8820 |
| TCTTCTTTTC CATCTGTAC AGGGGCTTTG CTATCTTGAT AGGCATCGCC TTGTTGAACG | 8880 |
| ATTTTGCAGG GTGTAAATGT AGAAGAAATA TAATCCATTA GGGAAGATGT AGCTGTAAAT | 8940 |
| CGAGCGTAAG GATTATTGTC TTGATGATCT GCATTTAAAA CAACTGTGAT GACTCTCATG | 9000 |
| CCTTTTTCGA CAGTAGTACC AACAAAAGAC TCTCCAGCCT TATCTGTTGT TCCTGTTTT | 9060 |
| AGCCCATCAA AACCACCAG GTAAGCAGGC ATACCTTCTA ACATGTAGTT GGTGGAAGTG | 9120 |
| ATTGTCATCC CAGCAAAAGT AGAAGAAGGT TTTTGGTGA TTTCTAAGAC TTGTGGGTAT | 9180 |
| TTTTTGATGA GGTGCGAGC AACGATAGCG ACATCATAAG CACTAAGCTT ATTTCTCTCA | 9240 |
| TCTTTTTTAG AACCTGGTA AATGTTATCC CCTAGAGTTT CATGTTAAG ACCTGTCGTA | 9300 |
| TTGACAACAG TGGCATCCTG AATCCCCAT TCCAAGAGTT TTGCCCCGAT CATATCGACG | 9360 |
| AAATCTTTTT CTGAGCCAGC AATTTTCTCA GCTAGGGCAA TAGCGGCGCT GTTGGCACTA | 9420 |
| GATACCAGAG TTGCTTCAAG CAACTCTTCG ACAGTATAAT TACGGGCCTC CATAGGAATA | 9480 |
| TTACTGGCTT CAGAATTGTT CGTCAATTGA TAAGGATAAT CAGAAATATC TACAGGAGTG | 9540 |
| GAGAGGGTAA TACTTCCGTT TTCCAAAGCT TCATAGACCA GATAAACAGT AATCAATTTT | 9600 |
| GTTATGGAAG CAATTCGAC AGGTGCGTT GCATCCTTCT CATAGAGAAT TTTACCAGTA | 9660 |
| TTTGCCCTCA CAGCAATCGC ATGTTTAGCG GCAATGGTAA AATCTTGAGC AACAGCAGTA | 9720 |
| GAAGCACCCC CTAAAAGAGA GACAGTTAAC AAAGTTAAAA ATATTTTTTT CATAGTAGTC | 9780 |
| TTATTTCTATC ATAAAGAAAA AAAATATTCT TGCTTTAATA ATTCACTGT TAAGCTTTTT | 9840 |
| GAAAATATGG TAAAATAAAG TAAGGGAGGT AACTCATGTT TCGTAGAAAT AAATTATTTT | 9900 |
| TTTGGACCAC AGAAATTTTA CTCTTAACCA TCATCTTTTA CCTATGGAGA CAGATGGGGT | 9960 |
| CTTTGATTAA CCCTTTTGTT AGCGTGCTTA ATACAATTAT GATTCATTT TTATTAGGGG | 10020 |
| GCTTTTTTTA TTATTGACA AACCTATTG TTACTTCTT AAATAAAGTC TGTAAACTCA | 10080 |

825

| | |
|--|-------|
| ATCGTTTGCT TGGTATTTTA ATTACCTTGT GTACTTTGGT CTGGGGAATG GTCATAGGTG | 10140 |
| TTGTCTATCT CTTACCTATT TTGATTAATC AGTTATCTAG TTTGATTATA TCTAGTCAAA | 10200 |
| CTATTTATAG TCGAGTACAA GACTTAATCA TAGACTTATC TAATTATCCT GCGCTCCAGA | 10260 |
| ATTTGGATGT AGAAGCTACA ATTCAGCAGT TAAACTTATC CTATGTTGAT ATTCTTCAAA | 10320 |
| ATATCCTAAA TAGCGTATCA AATAGTGTGG GGAGCGTCTT GTCAGCTCTT ATCAGTACTG | 10380 |
| TTTTGATTTT GATTATGACT CCAGTTTTTT TGGTTTATTT CTTATTAGAT GGACATAAAT | 10440 |
| TCTTGCCCAT GCTTGAAAGA ACGATTCTAA AGAGGGATCG CTTGCATATT GCAGGCTTAT | 10500 |
| TAAAGAATTT AAATGCGACG ATTGCTCGCT ATATTAGTGG AGTTTCGATT GACGCAATCA | 10560 |
| TTATAGGTTG TTTGGCTTAT ATTGGCTATA GTATTATTGG TTTAAAATAT GCTTTAGTTT | 10620 |
| TTGCCATTTT TTCTGGTGTA GCCAATTAA TTCCTTATGT GGGGCCAAGT ATTGGTTTGA | 10680 |
| TTCCTATGAT CATCGCAAAT ATATTCACTG ATCCCCATAG ACTGCTGATT GCAGTGATTT | 10740 |
| ATATGCTTGT TGTTCAGCAG GTAGATGGCA ATATCTTATA TCCTCGAATC GTAGGAAGTG | 10800 |
| TTATGAAGGT TCATCCAATC ACGATTTTAG TTTTACTTTT GTTGTCAAGC AATATCTATG | 10860 |
| GTGTAGTTGG AATGATTGTC GCAGTGCCAA CCTATTCTAT CTTGAAAGAA ATTTCTAAGT | 10920 |
| TCTTATCCCA TTTGTATGAA AATCATAAAA TAATGAAAGA ACGAGAAAGA GAATTAGCTA | 10980 |
| AGTAAAGTC AGGAGAACCC TGATTTTCT TTAAGTGAAG TGGCCTTTAG ATTAGAAGAC | 11040 |
| TGAAAAAATAG TTAAAGTCTT AACTAATTT TCACAGCTAA GAATAGTAGA AGTTAATCTG | 11100 |
| ATAAAAAATCG AAAAAACCAG TGGAATTCTG TGTGAGGGTA AGTTCCACTG GTTTTCATAG | 11160 |
| TCTATTAAAG TTCGAATGAA ACCTATTTAT AGTAGATTGA AACTAGAATA GTACACCTCT | 11220 |
| AATTCTAAAA CATTGTTAGA AATCGATTTG ACTGTCCTGA TCTATTCGTT CTATTTCTAT | 11280 |
| TTTATTTTAC TATATTTTGG TGCAATAAGT GAAAAGTAGT CCGAATAATA TAAGGATTGA | 11340 |
| TTTTATAGTT TTTAACTCA AATGAATTGA AATAAGAGA GTACGAAAAT TCTCATCTGA | 11400 |
| AAGTATTTTA GAATAATTCT CTTCTGGAAT TTCTTCAAAA CAGATAGCTT CATCTTAGGT | 11460 |
| ATGTGATTTT TTTTGCATT TTTGAGTTAG ATAAGGTATA ATGATTTTAT TGTCTTTTGG | 11520 |
| GGTCGTTACG GATTCGACAG GCATTATGAG GCATATTTTG CGACTCGTGT GCGGACGTAA | 11580 |
| ACGCTCAGTT AAATATAACT GCAAAAAATA ACACCTCTTA CGCTCTAGCT GCCTAAAAAC | 11640 |
| CAGCAGGCGT GACCCGATTT GGATTGCTCG TGTCAATGA CAGGTCTTAT TATTAGCGAG | 11700 |
| ATACGATTAA GCCTTGTCTA GCGGTTTGAT AAGAGATTGA TAGACTCGCA GTTTCTAGAC | 11760 |
| TTGAGTTATG TGTCGAGGGG CTGTTAAAAA AATACATAAC CTATGGTTGT AGACAAATAT | 11820 |

826

| | |
|---|-------|
| GTTGGCAGGT GTTGGACGT GGGTTCGACT CCCACCGGCT CCATTATTCC TTTGCATTCT | 11880 |
| TTTGCATTCC TTGGTAAAAC GTTGTTAAAT CAACGTTTTT TATTTTTATC TTTGGTATTC | 11940 |
| CTTTGCATTC TTTTGCTAAA AAGGGAGTCA CAAACAGACC CTATTTTAAA AAAGGATAGA | 12000 |
| AAAAAGGATA CAACATTGT CGCATCCTAA AAATAATCTT TTTTCGACGG AAGACATGGG | 12060 |
| ATTGGAACCC ACGCACGCTA TTACACGCCT ACCGCGTTTC CAACACGGCC TCTTAAGCCT | 12120 |
| CTTGAGTAAT CTCCAATAC TTA CTCAAAT AGTCTACCAT AAAGGCTCTT ATCTTGCAAT | 12180 |
| AAAAATTCTA GAAATAAGAA AAATGATAGA TTTTGAAAGA AAATGATAAA AAATGCTTGA | 12240 |
| CTTCGAAAGA AAGTATGATA GAATGAATAG TGTAACGAT AACAGGAGGT GATTCACTGT | 12300 |
| TAAAAACAGA ACGTAAACAA CTAATTTTAG AGGAGTTAAA TCAACATCAT GTAGTTTCTC | 12360 |
| TAGAAAAATT AGTTAGTTTG CTAGAAACGT CAGAATCAAC GGTTCGAAGA GACTTGATG | 12420 |
| AGTTGGAAGC GGAAAAAAG CTTCGTCGTG TGCATGGTGG AGCAGAACTC CCCTACTCCT | 12480 |
| TACAGGAAGA AGAAACCAT CAAGAAAAAT CTGTCAAAAA CCTTCAAGAA AAGAAATTGC | 12540 |
| TGGCTCAGAA AGCAGCCTCT CTCATTAAAG AAAAAGATGT CATCTTTATC GATGCTGGAA | 12600 |
| CAACAACGTC TTTTGTGATT CATGAATTGG TCAATAAGAA GTTACAGTT GTGACCAACT | 12660 |
| CCATTCACCA TGCCGCTCAG TTGGTTGAAA AGCAGAWTCC AACTGTCATG GTTGGAGGAA | 12720 |
| ACGTCAAGAC GGCGACAGAT GCTAGTATCG GGGGCGTTGC TCTTAACCAG ATTAACCAAT | 12780 |
| TGCACTTTGA CCGTGCCTTT ATCGGAATAA ATGGTGTTGA CGATGGCTAT TATACGACTC | 12840 |
| CTGATATGGA GGAGGGAGCT GTGAAAAGAG CTATTTTGGA GAATGCCAAG CAGACCTACG | 12900 |
| TCTTGGTGGA TTCGTCAAAA ATTGGACAAA CTTGCTTTGC CAAGGTAGCC CCACTCAAAC | 12960 |
| GCGCTATCGT TATCACTAGT CAAGGGCATG AGCTCTTGCA GGTATTAAAG GAGAAAACGG | 13020 |
| AGGTAATAGA AGTATGATTT ATACAGTCAC ACTCAATCCA TCCATTGACT ATATCGTTGC | 13080 |
| TTTGGACCAA GTCAAAGTTG GTAGTGTCAA TCGTATGGAC AGTGATGATA AGTTTGCTGG | 13140 |
| TGGGAAAGGA ATCAATGTCA GCCGTGTCTT GAAACGTTTG AATATACCAA ATACAGCGAC | 13200 |
| GGGATTTATC GGTGGCTTTA CTGGTAAATT TATCACAGAT ACTTTAGCAG AGGAAGAAAT | 13260 |
| CGAGACACGT TTTGTCCAGG TGGCAGAAGA TACTCGTATC AATGTTAAAA TCAAAGCAGA | 13320 |
| CCAAGAAACA GAAATCAACG GAACGGGTCC AACTGTTGAA TCGGTTACGC TAGAAGAATT | 13380 |
| GAAAGCTATT TTATCTAGTC TGACAGCAGA AGATACAGTT GTCTTTGCAG GTTCAAGTGC | 13440 |
| TAAAAATCTA GGCAATGTTA TCTATAAGGA TTTGATTTCC TTGACGCGCC AGACTGGTGC | 13500 |
| GCAAGTGGTC TGTGACTTTG AAGGACAGAC CTTAATTGAT AGTTTGGACT ACCAGCCTCT | 13560 |
| TCTTGTAATA CCAAACAATC ATGAACTTGG AGCGATTTTT GGGGTAAAC TCGAAAGTTT | 13620 |

| | |
|--|-------|
| AGATGAAATT GAGAAATACG CTCGTGAGTT ACTGGCTAAG GGTGCTCAAA ATGTTATTAT | 13680 |
| CTCTATGGCT GGTGATGGTG CCCTTCTTGT CACATCTGAG GGAGCTTACT TCGCTAAACC | 13740 |
| AATCAAAGGA ACAGTCAAAA ATTCAGTTGG AGCTGGTGAT TCTATGGTTG CTGGATTAC | 13800 |
| AGGTGAATTT GTCAAATCAA AAGACGTAGT AGAAGCCTTC AAATGGGGAG TGGCTTGCGG | 13860 |
| AACGGCAACT ACCTTCTCAG ATGACTTGGC AACGGCGGAA TTTATTAAAG AAACATATGG | 13920 |
| AAAAGTTGAG GTAGAAAAAC GATGAAATTT CAAGACCTAT TGAGAAAAGA TGTCATGTTG | 13980 |
| CTAGATTTCG AGGCAACTGA AAAACAGCT GTCATCGACG AGATGATTAA AAATTGACA | 14040 |
| GACCACGGTT ATGTAACAGA TTTTGAAACA TTAAAGAAG GAATTTTGGC GCGTGAAGCT | 14100 |
| TTGACTTCTA CTGGTTTGGG TGATGGAATC GCAATGCCTC ACAGCAAAAA CGCTGCTGTC | 14160 |
| AAAGAAGCGA CAGTCTATT TGCTAAGTCA AATAAGGGTG TTGACTACGA GAGCTTGGAT | 14220 |
| GGACAAGCAA CTGACCTCTT CTTCATGATT GCAGCTCCAG AAGGTGCCAA TGATACTCAC | 14280 |
| TTGGCAGCCT TGGCAGAATT GTCTCAATAC TTGATGAAAG ACGGTTTTCG AGACAAACTT | 14340 |
| CGTCAAGCAA CATCTGCAGA CCAAGTTATC GAACCTTTTG ACCAAGCTTC AGAAAAAAT | 14400 |
| GAGGAACTTG TTCAAGCACC TGCTAATGAC TCTGGTGACT TTATCGTAGC TGTTACAGCT | 14460 |
| TGTACAACAG GTATTGCCCA CACTTACATG GCCCAAGAAG CCCTTCAAAA AGTAGCTGCT | 14520 |
| GAAATGGGGG TTGGTATCAA GGTGAAACC AACGGTGCTA GCGGTGTTGG AAATCAACTA | 14580 |
| ACTGCAGAAG ATATCCGTAA GGCTAAAGCT ATTATCATTG CAGCAGACAA GGCCGTGAA | 14640 |
| ATGGATCGAT TTGATGAAA ACCATTGATC AATCGTCCAG TTGCTGACGG TATCCGTAAG | 14700 |
| ACAGAAGAGC TAATTAACCT GGCTCTTTCA GGAGATACTG AAGTCTACCG TGCCGCTAAT | 14760 |
| GGTGCCAAAG CTGCAACAGC CTCTAACGAA AAACAAAGCC TTGGTGGTGC CTTGTACAAA | 14820 |
| CACTTGATGA GTGGTGATC TCAAATGTTA CCATTGTTA TCGGTGGTGG TATCATGATT | 14880 |
| GCCCTTGCCCT TCTTGATTGA CGGTGCTTTG GGTGTTCCAA ATGAAAACCT TGGCAATCTT | 14940 |
| GGTTCTTACC ATGAGTTAGC TTCTATGTTT ATGAAAATTG GTGGAGCTGC CTTTGGTTTG | 15000 |
| ATGCTTCCAG TCTTTGCGGG TTATGTTGCC TACTCTATTG CTGAAAAACC GGGTTTGGTA | 15060 |
| GCAGGTTTCG TGGCTGGTGC TATTGCCAAA GAAGGTTTTC CCTTTGGTAA AATTCCTTAT | 15120 |
| GCCGCAGGTG GTGAAGCAAC TTCAACTCTT GCAGGTGTCT CATCTGGTTT CCTAGGTGCC | 15180 |
| CTTGTGGTG GATTATCGC AGGTGCCTTG GTTCTTGCCA TCAAGAAATA CGTTAAAGTT | 15240 |
| CCTCGTTCAC TCGAAGGTGC TAAATCAATC CTTCTATTGC CACTTCTTGG AACAATCTTG | 15300 |
| ACAGGATTTG TTATGCTAGC TGTGAATATC CCAATGGCTG CAATCAACAC TGCTATGAAT | 15360 |

828

| | |
|--|-------|
| GACTTCCTAG GCGGTCTGG AGGAGGTTCA GCTGTCCTTC TTGGTATCGT CCTTGGTGGA | 15420 |
| ATGATGGCTG TTGACATGGG TGGACCAGTT AATAAAGCAG CTTATGTCTT TGGTACAGGT | 15480 |
| ACGCTTGCGAG CAACTGTTTC TTCAGGTGGT TCTGTAGCCA TGGCAGCAGT TATGGCTGGA | 15540 |
| GGAATGGTGC CACCACTTGC AATCTTTGTC GCAACTCTTC TTTTCAAAGA TAAATTTACT | 15600 |
| AAGGAAGAAC GTAACCTCTGG TTTGACAAAC ATCATCATGG GCTTGTCAAT TATCACTGAG | 15660 |
| GGAGCGAATC CATTGGTGC CGCTGACCCA GCTCGTGCGA TTCCAAGCTT CATCCTTGGT | 15720 |
| TCAGCAGTAG CAGGTGGACT CGTTGGTCTT ACTGGTATCA AACTCATGGC GCCACACGGA | 15780 |
| GGAATCTTCG TTATCGCCCT TACTTCAAAT GCTCTCCTTT ACCTCGTTTC TGTCTTGGTA | 15840 |
| GGAGCAATCG TAAGTGGTGT GGTATTATGGT TACCTACGCA AACCACAAGC ATAAAAATA | 15900 |
| GAAAAATGAA AAGATTGGAC CGTTTGGTGC AGTCTTTTTT TCTTCCCGAA ATGCCTGTGA | 15960 |
| AATATGGTAT AATAGAAGAA TGGCAAACAA GAATACAAGT ACAACAAGAC GGAGACCGTC | 16020 |
| TAAAGCAGAA CTGGAAAGAA AAGAAGCGAT TCAACGAATG TTGATTTCGT TAGGAATTGC | 16080 |
| GATTTTATTG ATTTTCGCAG CCTTCAAATT AGGGGCTGCA GGTATAACCC TTTATAATTT | 16140 |
| AATTGCGTTG CTAGTGGGTA GCCTAGCTTA TCTGGCGATA TTCGGCCTAT TAATCTATCT | 16200 |
| CTTCTTTTTC AAGTGGATAC GAAAACAGGA AGGACTCTTA TCTGGCTTTT TCACCATATT | 16260 |
| TGCTGGCTTA CTCTTGATTT TTGAGGCCTA CTTGGTTTGG AAATATGGTT TGGACAAGTC | 16320 |
| CGTTCTAAAA GGGACCATGG CTCAGGTGTG GACAGATCTG ACTGGTTTTC GAACGACTAG | 16380 |
| CTTTGCTGGA GGGGGCTTGA TCGGGGTGCG TCTTTATATT CCAACAGCCT TTCTCTTTTC | 16440 |
| AAATATCGGA ACTTACTTTA TTGGTTCTAT CTTGATTTTA GTGGGTCTC TCCTAGTCAG | 16500 |
| CCCTTGGTCT GTTTACGATA TTGCTGAATT TTTCAGTAGA GGCTTTGCCA AATGGTGGA | 16560 |
| AGGGCACGAG CGTCGAAAAG AGGAACGCTT TGTCAAACAA GAAGAAAAAG CTCGCCAAAA | 16620 |
| GGCTGAGAAA GAGGCTAGAT TAGAACAAGA AGAGACTGAA AAAGCCTTAC TCGATTTGCC | 16680 |
| TCCTGTTGAT ATGGAAACGG GTGAAATTCT GACAGAGGAA GCTGTTCAAA ATCTTCCACC | 16740 |
| TATTCCAGAA GAAAAGTGGG TGGAACCAGA AATCATCCTG CCTCAAGCTG AACTTAAATT | 16800 |
| CCCTGAACAG GAAGATGACT CAGATGACGA AGATGTTTCTG GTCGATTTTT CAGCCAAAGA | 16860 |
| AGCCCTTGAA TACAACTTC CAAGCTTACA ACTCTTTGCA CCAGATAAAC CAAAAGATCA | 16920 |
| GTCTAAAGAG AAGAAAATTG TCAGAGAAAA TATCAAAATC TTAGAAGCAA CCTTTGCTAG | 16980 |
| CTTTGGTATT AAGGTAACAG TTGAACGGGC CGAAATTGGG CCATCAGTGA CCAAGTATGA | 17040 |
| AGTCAAGCCG GCTGTTGGTG TAAGGGTCAA CCGCATTTCC AATCTATCAG ATGACCTCGC | 17100 |
| TCTAGCCTTG GCTGCCAAAG ATGTCCGGAT TGAAGCACCA ATCCCTGGGA AATCCCTAAT | 17160 |

829

| | |
|---|-------|
| CGGAATTGAA GTGCCAACT CCGATATTGC CACTGTATCT TTCCGAGAAC TATGGGAACA | 17220 |
| ATCGCAAACG AAAGCAGAAA ATTTCTTGGA AATTCCCTTA GGGAAGGCTG TTAATGGAAC | 17280 |
| CGCAAGAGCT TTTGACCTTT CTAAATGCC CCACTTGCTA GTTGCAGGTT CAACGGGTTC | 17340 |
| AGGGAAGTCA GTAGCAGTTA ACGGCATTAT TGCTAGCATT CTCATGAAGG CGAGACCAGA | 17400 |
| TCAAGTTAAA TTTATGATGG TCGATCCCAA GATGGTTGAG TTATCTGTTT ACAATGATAT | 17460 |
| TCCCCACCTC TTGATTCCAG TCGTGACCAA TCCACGCAA GCCAGCAAGG CTCTGCAAAA | 17520 |
| GGTTGTGGAT GAAATGGAAA ACCGTTATGA ACTCTTTGCC AAGGTGGGAG TTCGGAATAT | 17580 |
| TGCAGGTTTT AATGCCAAGG TAGAAGAGTT CAATTCCCAG TCTGAGTACA AGCAAATTCC | 17640 |
| GCTACCATTC ATTGTCGTGA TTGTGGATGA GTTGGCTGAC CTCATGATGG TGGCCAGCAA | 17700 |
| GGAAGTGGAA GATGCTATCA TCCGTCCTGG GCAGAAGGCG CGTGCTGCAG GTATCCACAT | 17760 |
| GATTCTTGCA ACTCAGCGTC CATCTGTTGA TGTCATCTCT GGTTTGATTA AGGCCAATGT | 17820 |
| TCCATCTCGT GTAGCATTTG CGGTTTCATC AGGAACAGAC TCCCGTACGA TTTTGGATGA | 17880 |
| AAATGGAGCA GAAAACTTC TTGGTCGAGG AGACATGCTC TTAAACCGA TTGATGAAAA | 17940 |
| TCATCCAGTT CGTCTCCAAG GCTCCTTTAT CTCGGATGAC GATGTTGAGC GCATTGTGAA | 18000 |
| CTTCATCAAG ACTCAGGCAG ATGCAGACTA CGATGAGAGT TTTGATCCAG GTGAGGTTTC | 18060 |
| TGAAATGAA GGAGAATTTT CGGATGGAGA TGCTGGTGGT GATCCGCTTT TGAAGAAGC | 18120 |
| TAAGTCTTTG GTTATCGAAA CACAGAAAGC CAGTGCCTCT ATGATTCAGC GTCGTTTATC | 18180 |
| AGTTGGATTT AACCGTGCAG CCCGTCTCAT GGAAGAACTG GAGATAGCAG GTGTCATCGG | 18240 |
| TCCAGCTGAA GGTACCAAAC CTCGAAAAGT GTTACAACAA TAAAAAATA GCTTCTTTCC | 18300 |
| AAGTTTGGAG GGAAGCTATT TTAGTGGCTA TTGATTGCTT TTATTTTCTG AAGTTGGCGC | 18360 |
| ATTGGACTGT TTTTCGTTTT CAGTAGCAGG TTTACTTGAA GCAGGAGTAG AAGAGTCTG | 18420 |
| AGTTGCTGTT TTCTGATCTT CTTTTTCTC TTCCTTGACG CTAGATTTTG GTGTTTCCTC | 18480 |
| TTGCTGTGTT TTTTCTTGAC TAGTGTTAGT CTCTTTAGTT GGACTGGTGT TTTCTTTAGG | 18540 |
| GGATTCCTTT TGGATTCTT TGACAATGGT TGTCGTCTGG CTTGTCGTAG GTTCTTTTTT | 18600 |
| AATATTTTGT TTATTATCCA AGGCGTT | 18627 |

(2) INFORMATION FOR SEQ ID NO: 114:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2560 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 114:

| | |
|---|------|
| TAAAATACGT TACCTTGCTT CTGCACGTT AGCAGGTAAG TCATTGAAAT TTAAAGATCA | 60 |
| AGATATTACA ATTGAAGAAA CGACTGAAAC AGCTTTTGAA GGAGTTGATA TTGCTCTCTT | 120 |
| TTCAGCAGGT AGTTCTACAT CAGCTAAGTA TGCACCATAC GCAGTAAAAG CTGGCGTGGT | 180 |
| AGTAGTAGAT AATACATCTT ATTTCCGTCA AAATCCAGAT GTTCCTTTGG TTGTTCAGTA | 240 |
| GGTCAATGCT CATGCACCTG ATGCTCACAA CGGAATCATT GCCTGCCCTA ATTGTTCAAC | 300 |
| AATTCAAATG ATGGTGGCTC TTGAGCCGGT TCGCCAAAAA TGGGGCTTGG ACCGTATCAT | 360 |
| TGTTTCAACT TATCAAGCCG TTTCAGGTGC TGGTATGGGA GCAATTCTTG AGACACAACG | 420 |
| TGAACTTCGT GAAGTCTTGA ATGATGGTGT GAAACCACGT GATTTCATG CGGAAATCTT | 480 |
| GCCTTCAGGT GGTGACAAGA AACATTATCC TATCGCCTTT AACGCTCTTC CACAAATTGA | 540 |
| TGTTTTCACT GATAATGATT ACACGTACGA AGAGATGAAG ATGACCAAGG AACTAAGAA | 600 |
| AATTATGGAA GATGATAGCA TTGCAGTATC TGCAACATGT GTCCGTATTC CAGTCTTGTC | 660 |
| AGCTCACTCT GAGTCTGTTT ATATCGAAAC AAAAGAAGTG GCTCCAATCG AAGAAGTAAA | 720 |
| AGCAGCTATC GCAGCCTTCC CAGGTGCTGT TCTTGAAGAT GATGTAGCTC ATCAAATCTA | 780 |
| TCCTCAAGCT ATCAATGCAG TTGGTTCGCG TGATACCTTT GTTGGTCGTA TCCGTAAAGA | 840 |
| CTTGGATGCA GAAAAAGGAA TTCACATGTG GGTGTGTTCA GATAACCTTC TCAAAGGTGC | 900 |
| TGCTTGGAAC TCAGTTCAGA TTGCTGAAAC TCTTCATGAA CGTGGATTGG TTCGTCCAAC | 960 |
| AGCCGAATTG AAATTTGAAT TAAAATAGTC ATATCGTTTA GGAGTTCAGA TGAATCCTT | 1020 |
| CTTTGAAATA GAGAGGTGTT TTCGTGCTT ATCAAGATTT AAAAAAATGT AAAATCATTA | 1080 |
| CAGCCTTTAT TACCCCTTC CATGAGGATG GTTCCATPAA CTTTGATGCT ATTCCAGCCT | 1140 |
| TGATTGAGCA TTTATTGGCC CATCATACGG ATGGAATTCT TCTCGCAGGA ACGACTGCTG | 1200 |
| AGAGTCCAAC TTTGACCCAC GATGAGGAGT TGGAGTTGTT TGCGGCTGTA CAAAAGGTTG | 1260 |
| TCAATGGACG CGTTCCTTTG ATTGCGGGTG TAGGTACTAA TGATACGCGT GACTCTATTG | 1320 |
| AGTTTGTCAA AGAAGTAGCG GAATTGGTG GTTTCGCAGC TGGGCTTGCT ATTGTTTCCTT | 1380 |
| ACTACAACAA ACCTTCTCAA GAAGGGATGT ATCAGCACTT TAAGACTATT GCAGATGCTT | 1440 |
| CTGACCTACC AATTATTATC TATAACATTC CAGGGCGTGT AGTTGTCGAA TTGACTCCAG | 1500 |
| AAACCATGCT TCGCTTGGCT GACCATCCAA ATATTATCGG TGTCAAAGAA TGTACTAGCT | 1560 |
| TGGCTAATAT GGCTTACTTG ATTGAGCACA AGCCTGAAGA GTTCTTGATT TATACAGGTG | 1620 |
| AGGATGGAGA TGCTTTCCAT GCCATGAACC TTGGGGCGGA TGGGGTTATT TCTGTTGCCT | 1680 |

831

| | |
|---|------|
| CTCATACAAA TGGGGATGAA ATGCACGAGA TGTTTACTGC GATTGCAGAA AGCGATATGA | 1740 |
| AGAAAGCCGC AGCAATTCAG CGTAAATTCA TTCCTAAGGT TAATGCTCTC TTCTCTTATC | 1800 |
| CAAGTCCTGC TCCAGTTAAG GCAATTCTTA ACTATATGGG ATTTGAAGCT GGACCCACTC | 1860 |
| GTCTACCTCT TGTTCAGCA CCAGAAGAAG ATGCCAAACG CATTATCAAG GTTGTCGTAG | 1920 |
| ATGGCGACTA CGAAGCAACT AAGGCAACTG TAACAGGGGT CTTAAGACCA GATTACTAAT | 1980 |
| AAAGACAATA AAATCCGGCT CTTTGTCAAC TGTAGTGGGT TGAAGTCAGC TAAGCTCGAG | 2040 |
| AAAGACAAA TTTTGTCTTT TCTTTTTTGA TATTCAGAGC GATAAAAATC CGTTTTTTGA | 2100 |
| AGTTTTCAAA GTTCCGAAAA CCAAAGGCAT TGCCTTGAT AAGTTTGATG AGATTATTGG | 2160 |
| TCGCTTCCAA TTTGGCGTTT GAATAGGGTA GTTGAAGGGT GTTGACGATT TTCTTTTTGT | 2220 |
| CCTTTAGAAA GGTTTTAAAG ACAGTCTGAA AAATAGGATG AACCTGCTTC AGATTGTCTT | 2280 |
| CAATGAGTCC GAAAAATTTT TCCGGTTCCT TATTCTGAAA GTGAAACAGC AAGAGTTGAT | 2340 |
| AGAGCTGATA GTGATGTTT AAGTTTGTG AATAGCTCAA AAGCTTGTTT AAAATCTCTT | 2400 |
| TATTGGTTAA GTGCATACGA AAAGTAGGAC GATAAAATCG CTTATCACTC AGTTTACGGC | 2460 |
| TATCCTGTTG AATGAGTTTC CAGTAGCGCT TGATAGCCTT GTATTGCGGA TTTTCGATGA | 2520 |
| AAGTATTCA TGATTGGAC ACGCACACGA CTCATAGCAC | 2560 |

(2) INFORMATION FOR SEQ ID NO: 115:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 11303 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 115:

| | |
|---|-----|
| TATTGGATTT CCCTTGCAAT CAGTTTATGG GACAAGCACC CGGCAGCGCA GAGGAAATCA | 60 |
| ACGCCTTCTG TAGCCTACAT TTTCAAACCA CCTTCCCACG TTTTGCCAAG ATTAAGGTCA | 120 |
| ACGGTAAGGA AGCAGACCCT CTCTATGTCT GGTACAAGA CCAGAAATCC GGCCCACTAG | 180 |
| GAAAACGAGT CGAATGGAAT TTCGCTAAGT TTCTCATCGG TCGAGATGGG CAAGTCTTTG | 240 |
| AACGCTTTTC TTCAAAAACA GACCCAAAAC AAATTGAAGA GCGGATACAA ACTCTACTAT | 300 |
| AATTCACAAT CTCATATGA TTAGGTTTCC TTAAACCTGA TGAATAGTGA GATTTTTTGA | 360 |
| TGGGCTTTGA CTAAATAGA AAAACACCCC ATGATATGAA ACATGAAGTG TTGTAAAGTC | 420 |
| TATGTTGTAG GTGCTTATTT CACAATTTC AATGTGACCAG TGATAACGAA TACCATACAG | 480 |

832

| | |
|--|------|
| AATCTTCATA TACACTAAAC AAATGACTTT CTAATTATTT CAATTAGTTT TGGCTAGTAA | 540 |
| ATATCATTTTC CAACAAACGC CCTCTCAATT CCTTATCCTG ATGATGCAAG ATATTCATTA | 600 |
| AGTCATGAGA GTTTTTTCGCA TTGATGAATT GATTAAACAA TCTATCTTTT AATTCATATG | 660 |
| GAAGAGAAGC TGTCTTTAGT AGTCTAAAAA CTTGTCATT TAAAGATGTC CTTTTATTAT | 720 |
| CTTTCCATTC AAATTTAGCT GTATCATTTCT TATTTGGCAA TTCAATTATA GACACATTCG | 780 |
| TTCTTTTAAA ATGAATTCTA TGTCTTCTAT TGCTTGGAA CATACTAGAA TCTCCTTGTA | 840 |
| ATGCTAACTC TACCATTCCC ATTTCCCAAT CGATTGATAA TCTTGTTTTA TATCTTTGAC | 900 |
| CATTTTGATC TTCAAGCATT TCAAAAGAAT GTTGTTTTCC TGGGAATACA TACCAATCTA | 960 |
| CAACTTCAGG TAAATCAACA CCCATACCTA TCTCAGAACC AACCAAGGGA ATGATTGCAC | 1020 |
| CACTTTTTGC AAACACAGGC GTAGTCGAGA TGTCCTTATA AACACTTAAC TTCACACCAC | 1080 |
| CTGTGTATTT TTTCTCTGAA AAGAAGTCAT ACCATTCACT TTCAGGGAAC CATACATCTA | 1140 |
| CTTTTGCAGA TTGGAATGTC AAATCCATCT TTTCTACAAT GGGAGCCACC ATCAGTTCTG | 1200 |
| TTCCAAAAAA GTATTGGTTT GGAACATTAT AGCTCTCATC ATTCTCTGGA TAGAAATAAT | 1260 |
| AGATTGGACT GATTAATGGG GCACCTTCCT CATGTGTCTG TACATTCATG GTATATAGAT | 1320 |
| AGGGAATCAT CTGATGTCTC AAACGAAGGT ATTTCTTCAT AATCTTAGAT GTTGTTCCTG | 1380 |
| AAAAAAACCA AGGTTCTTTA CTATTAAAAG GACTTCTAGA ACTATGTAAT CGAGTAATCG | 1440 |
| GACTAAAAAC ACCAACTGT AGCCATCTAG TTTGTAGCTC TTCGTCATAA TCCCCAACA | 1500 |
| TATGTCCACC GATATCATGA CTCCACCAAC TATAACCGAT ATTAGATGCT GTCGCTGTAA | 1560 |
| AATAGGGTTG AAATCTTAAG GAATTCCAAC TAATAATAGT ATCCCCTGAA AAACCAACAG | 1620 |
| GGTAGCGGTG ACTACCAGGA CCTGCATATC TTGATAAAAT CAAACCACCT TCTGCATTTT | 1680 |
| TACAACTATC CTGATAGTGA TAATGGTTTA AAAGCCAAAG TGGATCTAGC ATACCTTGTC | 1740 |
| TCCCTTGTTG CCAGTCAATC CACCAAAAAT CTAATCCCTG CTTTCTAGT TCATAATGAA | 1800 |
| CATCTTTAAA GTAGGCTTCC CTAAAAGAGG GATTAAAAAA ATCAAAAATA GCAGGTTCTT | 1860 |
| CTAGTTCTAC ATTTAACCCC AACCGTTTTCG CGATTGAGG ATAAGCTTCT TCATAAGCCC | 1920 |
| GTATCCCATC AGCAGGATGG ACATTTAAGG AGAGTTTTCG CTTTCTATCA TGAAGTTGTT | 1980 |
| GCAATAACTG TTCTGGATTT GGTATTAAGT TTCTATTCCT ACTATATCCT GTCCAGCCAC | 2040 |
| TTCCAAAGCG AGCTGGAATG TCAGTTATAT GCCAATCCAT ATCTAACACA CCGATAGATA | 2100 |
| ATGGAATTTT CTCTGTTTCA AATCTGTCTA TTAAATCCAA GTATTCATCC GACGTATAAG | 2160 |
| GCCAATATCT ACTCCACCAA TTGCCTAAAG CATATCTTGG CAACAAGGGT GTTGAACCAG | 2220 |
| TCAAATGGTA AAAATCTCTG ATTGCTCCTC TATAATCATG CCCATAGGCA AAGAAATACA | 2280 |

833

| | |
|---|------|
| GGTCAATTTG ATTTTCTCTC TCAATATAAC CAGATTGTTC ATCCCAAATA AATCCTTGAG | 2340 |
| AATCATCCAA TAAGGCTATA CCATTTCGGC TAATAATTCC ATCTTCTAAC GAGATTGCTC | 2400 |
| CATCTGCCTT ATCCAGAGTC CGAGCTGTTC CTTTAAACGT TTCAATAGAT TCACCAAAAT | 2460 |
| ACCAGCGACT ACCATATACG GCAAAATTTC CTTTAAATTC TATAAATAAA TTTTCGGCGT | 2520 |
| TAAATTCTCC TTTATTAAAG TGCAGATGAA AATAGTCCGT CATAATATCT AGTACGTTTG | 2580 |
| ATGTCTCGAT ATAATCTAAC GAAATTGGC CAAATCTCT ATTATAGATA AGTTGTGTCG | 2640 |
| TTCTATCCTC AAAACTTCCA GTTTGAGAGT ATTCTAACCT TACTAGCTTG TCTGTTAATA | 2700 |
| CAGAGATTCG ATAAACTCT CCCTTAAAAA TTTTCAATTT GTTTTCCTCC TTTTATGGTA | 2760 |
| GCATAAAAC AGAACGCACC ATTTTGTATG CGTTTTTCAT TATTCTGAAT GCAATGTTCT | 2820 |
| ATCTGTTATA TCTATGACAA ATAATAGTCA ATTGAAAAA TGCAGTGGAC AAAATATCTT | 2880 |
| TTAACAAACC AAGAGTTTAT TAAAGAGTTA TCACTTTTCA ACTTTTCTAA GCTTATGCAG | 2940 |
| TTGTGAAACA AACTACTTTT AAATATTAA CTAAGATAGG ATTGATAAAT AATTTCAAAC | 3000 |
| TCTTACTAGC AATCATACGA TATTCAAGCT CACGTGCTTT TTTCTTCCT GCTTATTCT | 3060 |
| TAGAACTGAA GAACCCGGAT CGGTATATAA ATTATCCGA TCAACATAGT CATAAGATTC | 3120 |
| ATAACAGTTG CGCTTCATTA AGTCATCCCC AGAGCAAGAG CTTCATCTCG TAATTTTCA | 3180 |
| ACATCACTAA CCGTAGGTCG CCATCCTTCA ATCATATTTG TACTTAAAGC ATACCAACA | 3240 |
| CTCTTAAAA CGGATCGGTT TTCAAAGCT ATTCCCATGA TTGTCATCTT TTCTTTATCT | 3300 |
| ATATCTAAGG ACATATGCTA CCTCCTTTAG ATACATTATA CCATGTTTCT CTGTAGCTTT | 3360 |
| TAAAAATTTT ATTTTGTTTG TCATATCTAA GTTTTCAGCA CGCTTATCCT ATTTTATAAG | 3420 |
| CCTCAAACCC AAATATAAAA CGCATTCCTT TTGCTTTTTT ACTATTGTAT CGTATTCTAC | 3480 |
| GATAACATAC TTTACTTTAT TGTTTTTTTA AATAACAGCA GTTCCCTGTT TATCAACTAT | 3540 |
| TCGAACTACT TTCTATTTTG CTTCATACCC TACATAGCGA AAAAATATGA AAAAGCAGAG | 3600 |
| AAGAATATCT TAAAAAGACC TCTTCACTGC TAATATTAAC ACTCATTATT TAAACTATAT | 3660 |
| GGATFCTATC ATCGAGTATA CTTTTTACT TATTAGATAC CTTGCTCTTC TTTCACCAAT | 3720 |
| TTTTGATCAT ATACACGGAT GAATGGAAGA TAGACTAGGA ATGCTGCAAA TGCACATACT | 3780 |
| AGAGCAACTA ATACAGCTCG AAGATCTGCT GTCCCTAAGA AAGCTCCAAT CCCTACTGGA | 3840 |
| GTTGGCCATG GAACCTGTGC GATAATTGGC TTAATAAAGT TTAGAGAAAT CGCTACGTAA | 3900 |
| TAAATAGTAG CAGTAACCAT TGGTGCTAAA ATAAATGGTA TAGCCAAGGC TGGATTATAG | 3960 |
| ATAATAGGTA ATCCAAAAAT TAATGGTTCA TTAATATTAA ATAAGGCTGG AACTACAGAT | 4020 |

834

| | |
|--|------|
| GCTCGTCCTA TTGCTTTAAG CTGTTTCAGAT TTAGAGGCAA AAGCAATATA TAAACATAGT | 4080 |
| CCTAAAGTTG CACCAGAACC ACCTGCAATT ACAACATAT TAGAAAATTC ACCTGCAACA | 4140 |
| GCGAAGTGCC CGCCAGCAGC ATTTTCAGCC ATGTTAGCAA GAGCAATTGG ACTAACAAAT | 4200 |
| GCAAAAACAA TGTTGCACC GTGGATACCT ACAATCCAAA GTAGTTGAGT CAATAGATAA | 4260 |
| ATAATCATTA AACCAATCCA CGAATTAGTC AGATTGGATA CAAAACCAA TGGAATTGCA | 4320 |
| ATGACTTTAA AAATATCTGT TCCCATTGCT ACAAGAAGAC CGTTGATAAA GATAACAACA | 4380 |
| AATGCAACAA CAAATCCCGG AACCAAAGCG GTAAATCCAC GAGAACTCC TTCTGGAACA | 4440 |
| GCTTCAGGCA TTTTAATAAC CCAATTATGT TTAACACACA TACGATAAAT AAGAACAGTC | 4500 |
| ACAATTGCCA TAATGATTGC GGTAAAAATC CCTGTTGTCC CAAAACGTGC GACTACATTT | 4560 |
| CCCATTGCCC ATCCATCTGC AATTACTGCA CCTTCTTTTA GACTTGTCAC AGTCTTCATC | 4620 |
| ATTCCACCAT CAAAAATGAT TTGCGGTACT GTCATGACAA AAGCCATCAA GGCAAGCAAG | 4680 |
| GCACCATTAA GAGGATTTCAT ATTGAGTTCT TCTTCTCTG CATAAATTTT TGTCAATTCA | 4740 |
| TATGCAAGTG ATAGAACGAA ATAAAGAGAT AGAGAACCCA TAGTCGCATA GTTTGCAACC | 4800 |
| ATGTAAGTG ATGTGAATTT ATCAAATGAA GCAGAGAAAA TATCTGCCAC AATTGGCCAA | 4860 |
| AATGAGAAAAG CTGTGGCAA AATACTGAAT ACCAAAAACA TTGATCCTAC AATAGTAAAT | 4920 |
| GGTACAGCAG CCATACCTGC AGCCGTGATA GCACGTACTA CTTTAACTG AGCAAGTTTG | 4980 |
| CCCATTGGTC CCATAACATG GTTTTCAAGA AAACCAAACA ACCCGTTTG TTGATCCATA | 5040 |
| AATAGACCTC CTTAATAAAA CATAATAATT TTTACTTTCT AAAGACTAGT TTCAAATACA | 5100 |
| AATTATACTA GATCAGGATT ATAACTAAG TGAGTTCTTT TCCAATTGGA CAAATGTGTG | 5160 |
| ATAAGCCTTA TCTGTTCTGT TATAAATTTT TTTAATTCTT CTAATGTCTA ACAAACTCAG | 5220 |
| AACTAAACCT AATAGAAGAA CTACAAAAC AAATAAACGT GCTACTTGGT TATTTTCAAA | 5280 |
| AATCGGAAAA AGATTCTTAA ACCAACTTGT CCAAGTTAAA ACAAGTAATC CTATTGAAAT | 5340 |
| AAGCATTTGT ATTCTAACAA ACATTAGTGT TATTCCTAAC TTTTCTTTCC TATTTCCATA | 5400 |
| AAGTTTAAAT TGTTCACAG TTGCTAAAAT AGAAAATACT ATGAGCATAA TGCGGAAAAT | 5460 |
| AATAATAGGC GAGGGACTAA TAACTGACT CAAAAGCCAA TAAATATTCC CAAAAAGAA | 5520 |
| GAGTGCTATT GAATAACGTA GAAGAAGATA TCGATTGAAA AAAGTATTAG TTAGAGCCAT | 5580 |
| CTCTCGACGT TGTGTTCAA TCTTTGTGCG TTCTTTTTTA TCCATATCAT TTCTCTCTTA | 5640 |
| TATAACAACA CATATTTAGT TAACTTTCTT ATAAAGAGCT AACATTTCCCT TTGCTACTTC | 5700 |
| TAATAATGTC ATAGTGGTCA TTAAATGATC TTGAGCATGT ACCATGATAA TTTCAATTTT | 5760 |
| AATTTCCACT CCACTTGCGT ATTCTTGCAA GAGTTTGGTT TGTGCATGAT GCGCTTCAAG | 5820 |

835

| | | | | | | |
|------------|------------|------------|-------------|------------|-------------|------|
| AATTATCTCA | TTTGATTGAT | TTAATTTACT | TTCTGCATCA | TCAAAACTAC | CTTCTCTCAT | 5880 |
| TTTTGCAAAT | GCTTCATGTA | TTTCTGACCT | TGCATTTCCC | GAATGCAGGA | TAATTTCAAA | 5940 |
| TGCTGCAACC | TGCAGTTCCT | CTTGATTCAT | ATAAACCTCC | TATTTTATCT | TCTCAAATAT | 6000 |
| GTTAATAAAA | TCTTCAAAGT | TATTGCAAGA | TATTAGCTGA | TTTTGCAATT | CATCATTCCTC | 6060 |
| TGTCAGAGAG | ACTATCTTTT | TAGTCACAGT | TGCCAAACCT | TCGTTCCCAT | ATATTGATGG | 6120 |
| AGATAGAAGA | AATACTAGCT | GGACATGTGA | ACTTTGATTA | TCCCAGAGTA | ACGAATCTTT | 6180 |
| ACAAATTGCA | ACCGAAACCT | TTCCCTCTGT | ACCAAAGGC | TGAATAGGAT | GCGGAACTGC | 6240 |
| AATTTTTTCA | GAAAAACAA | CTGAACCTAA | TTCTTCGCGC | TGTTTAATTC | CATAAAGTAA | 6300 |
| AGATTGTTCA | AACTCATTTG | ATTCACCAAC | AGATAAACTC | TCAACCATCT | TTTCAAGTAA | 6360 |
| ATTTACCTTG | TCTGATTCAG | TACATATTAA | AAAGTTTTCT | TTACTAAAAT | ACTGTCTAAA | 6420 |
| GCCGTTGTTT | TCAAATTTGT | TAATCTTTGA | TGATTGTACA | TAACTAGAAA | CTTGCATCTA | 6480 |
| ATCCATAGCT | TTTCTAATCA | TTTCCATCTC | ATCACTCTTA | AGAAACACAC | TAACCTTTAA | 6540 |
| AACTGGGATT | TGAAAATATA | GATTTGATAA | ATCAATAGCT | GACACTATAA | AATCTATTCC | 6600 |
| TTTAAGTTTT | TCTTGATTCA | ATTCATAGTA | GCCTATTACA | TCAACAACTT | CTACTCGCTT | 6660 |
| CCCAAACCTC | GTTTCCAAC | GATTTCTTAA | CATTGGGGCT | GCACCAAATC | CTGTTGCACA | 6720 |
| AATAGCAAGA | ATATTAAACT | TAGTACTCTC | TTTGCTACGT | TCCATAGCAG | CTAAAAAGTG | 6780 |
| AAGACTTACA | TATGCTACTT | CATCATCTGA | TATTGTCCAC | TCCAAGAACT | TGTCCATATT | 6840 |
| TGCAAGAATT | TCTCTAGTCA | TAAAGAATAT | ATCACTATAA | TTCTGTTTAA | TTTCATCTAC | 6900 |
| CAAAGGGTTA | TTTAAGGTAA | TCCGGCTTTC | TAAACGTACT | TGTAGTGTCA | TTAGATGAGT | 6960 |
| TATCAATCCT | TCAATTAGTT | GGAAATCTGA | AGAAAAGTTA | TACATATCAT | CTAATCCTAA | 7020 |
| ATTCTGAAAT | GTTTTAAATA | AAGATTTTTT | TAAAACTTCT | TCAGAAATAT | TCTTCTGATT | 7080 |
| TTTTTGACAT | TGTTGACTCT | TAGCTAACAA | ATGCAAAGTA | ATGTAGTCTA | TTTCTGGAAC | 7140 |
| TGGAAATTCC | TGATTGTGTA | CTTCTCTTAC | TTTAGAAAGA | ATTCTTTGGG | CAACCTTTCT | 7200 |
| CTCTATTGCA | TCATCAGTCA | TCTGACAGTC | TATATTTTTT | ATTTCAAATC | CGGATTTTAA | 7260 |
| ACGAATCACA | GACAATGCTA | TGTGAACTAC | TAAATTCCTGT | AGTACAAAAT | CAGATAGTTT | 7320 |
| TAGGTTGGCC | TCTTGGCATT | CATCCAAAAC | AATTCCTAGCA | AATTCCTCTA | ATGGAACAGT | 7380 |
| TTGATCAAAA | AAGTTAAATT | TTACATAGCA | ATGTATTGTT | TTAAAAAATT | GATTCTCTAG | 7440 |
| GAAATAATTT | ATGATAAAAC | GTCGTTTATC | ACGTTCCCTCG | CCTGAGACAT | AAACTCCTTT | 7500 |
| ATTCGCCCTA | CTCTCAATGG | ACAAATTATA | CTCTGATAAC | ATCACTCGTA | TCTTTCTGAA | 7560 |

836

| | |
|--|------|
| ATCATGAGAT AATGTTGAAC GACTAACGTA AAGTTCATCA GCTAAATCAT CAAAAAGAAC | 7620 |
| TGGAACCTGC TCAATAATA ATTTATTTAA GATAAATACT AAACGATCAT CACCTTTTGA | 7680 |
| AACCGCAGTT TTCGTATAGT CTTCTTCCAG TTCATAAGTT TGTCTAAACT CCTGGTAAGC | 7740 |
| GCCTTGATTC TCAAAAAATA TTTGATACCC TTGACCTTGT TTTGAAATCA ACCGGACTCC | 7800 |
| TTGAATAATC ATTGTCTTCT CAATTAATTT CAGTACATTA CGGACAGTTC TATCTGAACA | 7860 |
| GGATAAATAT TCTGCCAGTT CTTTGCTTGT AACAAAACGT TCCTTATTTT TTATTAAAAA | 7920 |
| TTGAAGGATA TCTTTCTCTT TAATGTTTAA CACATTCATT CCTCCTAAA ACGTATGTTT | 7980 |
| TCATATATTG AAGCATATTA TACACTTAAA TCAGTTTATA TCAAACCTCAA AACAAATTAT | 8040 |
| CTTAACCTAA ATATTATATG ACATTTTCATG TGTTCATCAA ATATTCTCAA GAATCAAATT | 8100 |
| AGCCATTTTT TCAATTCCCA TTGGAATAGG AATATAGGCT TGAGGAGGTA TTTGTACAAC | 8160 |
| TGTTTTTCCT GCTTTAGAAC CAGCCTCTTC AAATTGCTTA AAGTACATTT TTGTTTGAGG | 8220 |
| ACTGACAAGA TACAAATCAA AAGCTGCTGC TGCATAGCT TTCCCTCCTT CAGTAGCACT | 8280 |
| AATAGCATCA ACTACAATAT CTTTCCCTTT TCCTTTTAGA AACTCTGTTG TTTTCTGTGC | 8340 |
| CATAAGTGAT GAAGACATTC CTGCTGCACA AATAATTAAA GCTTTTGCCA TAATATTTTC | 8400 |
| TCCTTTTCTT AAATCCAATC AAAGCTGTGC TAAGTTGGCT TATTTGTTAT CTATTTTTAT | 8460 |
| TATAAAATAA AGCGTTTCCA ATGACAATTC CCTCATTTTC CTAAATGATA TGGAAAAAAA | 8520 |
| TTATTTATAC TTCAATTTAT AAAATAAAAT TATTCCTGAG AGTAGAAATG AAACACTATT | 8580 |
| TGCTAAAATC AAAGGCAAGT CTCCTATACG AATACCATGA GCAAGCCACA ATGCAATACC | 8640 |
| AATAACTTGC ATAACATACA TACCTAGAGC AATAGATCCT GTGTCCTTTG TCTPAACTAC | 8700 |
| ACGAAAAACT TGTGGTAAAA ATGCAAATGT TGTTAAAATT GCTGCAATAC TTCCAATCAT | 8760 |
| ATGTCACCTC AATATGCTAA ACAAACTGAG AATAATCTCA GTTTGTTTAT ACTATTCTAC | 8820 |
| TGATTACCGG TTAGATGAAA TAACTTCCTT ATACCAGCCA AAAGATTTTT TCGGGGAACG | 8880 |
| ATTATAACTT CCCTTCCCAT TATCATCTTT ATCTACATAA ATAAAGCCAT AACGTTTCCG | 8940 |
| CATTTACCGG GTACCAGCTG AAACCAAATC AATACATCCC CATGGAGTAT AACCCATTAA | 9000 |
| ATCAACACCA TCTTCAACTA CAGCCTTTTT CATTTACGA ATATGGGCAC CTAGATATTC | 9060 |
| AATTCATATA TCATCATGTA CCATACCATC TGCTGCAACT TGATCTATAG CTCCAAAACC | 9120 |
| ATTTTCAACA ATAAAGAGTG GTAAGTGATA GTGGTCTGTA AACCAATTTA ACGCATAACG | 9180 |
| CAAACCTTCT GGATCAATTT GCCACTCCCA TTCAGAAGCC TTAACATAAT TATTTTTCAC | 9240 |
| TAAATCTTCT GTTCAAGAT AATCAAAATA AGGATTATTT TCACGATGAG AGTCGATAGC | 9300 |
| AAAGGACATA TAGTAACTGA AACCAATGTA ATCTACAGTC CCACCAAGTA AATCTTCTTT | 9360 |

837

| | |
|--|-------|
| ATCCTGGGCA GTAAATCAA CTGAAATACC TTTTCGTTCC CAATACTTGA AAATATGCTC | 9420 |
| AGGATATTTA CCTAAAACAT GCACATCAGC AAAATAATAA CGCTTCTGCA TAGCTTTTCAT | 9480 |
| TGCCATTAAAG ATATCCTTAG GATTGCAAGT AACTGGATAA ATTGGACACA TCGCAATCAT | 9540 |
| ACAACCTATT TGAAATCTG GATTAATCTC ATGACCAATT TTTACAGCTC GTGCAGAAGC | 9600 |
| AACTAATTCG TAATGTGCTG CTTGATACAT AATTGCTTCT CTATTATCAC CTTCTCATA | 9660 |
| TACAATACCT GAGTTAGTAA ATGGTGCAA ATCTTCCTGA TAATTCGCTT GATTATTGAT | 9720 |
| TTCAATTGAAA GTCATCCAAT ATTTAACCTT ATCTTTGTAA CGTTTAAATA CGACTTCTGC | 9780 |
| AAAACGAGCA AAGAAATCAA TCAATTTCCT ATTTTTCCTT CCACCATATT CGGTCACTAA | 9840 |
| GTGATAAGGC ATTTCAAAAT GAGATAGAGT GATGACAGGT TCAATACCAT TCTTTAAGCA | 9900 |
| TTTCATCAAAA AGATTATCAT AAAACTGTAA TCCTTCTTCA TTCGGCTCTA ACTCATCACC | 9960 |
| TTTTTGAAAG ATACGTGTCC ATGCAATAGA GGTACGGAAG CACTTGAATC CCATTTTCAGC | 10020 |
| AAAAAGTGCT ATATCTTCTT TATAACGGTG ATAAAAATCT ATCGCCTCAT GATTTGGATA | 10080 |
| ATATTTACCC TCTAAAACTC CCAAAGTAAT TTCACGAGCT ACTCCATGAC GACCAGCAGT | 10140 |
| CATAACATCA GCAACACTAA TTCCCTTGCC ACCTTCTTGC CATCCACCTT CAAGTTGATG | 10200 |
| AGCAGCAACA GCACCACCCC ATAAAAATCC ATCTTTAAAA GTAGTCATCT TTTTTCCTCC | 10260 |
| TGACTTTGAT ACTCTTATTA TAAACCTTAA ACCAAAAGAT GAAAACGCAT TCTTTTTCCT | 10320 |
| TATTGTAAAG GAAAGAAGTA ATTTTAAATG GAAATAGAAC AATATCTTCT TGTATTCTCG | 10380 |
| TAATGATATC TTTACGATT TCAATACTTT CAAACTACAA AAACCTCTAC AATAATTCTA | 10440 |
| ATTCCTGTG TCTATAAAG ACTTATCGCT TTCTGGCATC CCAGAATCAT CTTCTATATA | 10500 |
| ACGTTCAACT TGCATCTGCA AGTGATATTT TTTTCTTAAA TCTAAGATTT TCTGCATTGT | 10560 |
| CTTTGATTGA TAATGTTTAT CTAAAGTTTC TTGATTTATC CACTGATCAA TAAGGAGAAT | 10620 |
| AGTTCCCTCT TTTTCAATTG GTAAAAATA TTCGTATTTT AAGTTACCTT TTTGATTTCT | 10680 |
| AATTTCTTTA ACAAGGCCAC TATCAAGCAT TTCTCTTGCA AACTTTATTG CACTATCTCC | 10740 |
| ATCACCTTTA TAATATACAT GAATAGTCAA TGTATCTTTA TATCCTCCAA AATCATCCTT | 10800 |
| CAATTTTAAA AAAACAAGTT TAGATGAGGA TCTAAACTTG TTTTATATGA ACTAATTATC | 10860 |
| TAACGTTTCG CCATTACTTT CAATCACTTC TTTATACCAA TAAATGATT TTTTCTTATA | 10920 |
| GCGATTTATA GTCAATTGAA ACAAGAGCAG GACAAAAGAG CCTCATAAAA GGTATTGCAA | 10980 |
| CTTGGTAATA CCTTTTGTAG GTGCTTTTGT ATATGAGCCC ATGTTTCTC AATAGGATTG | 11040 |
| TACTCAGGTG AGTAGGGAGG AAGAGGTAAA AGTTTATACC CAAACTCTTC ACACAAGAGT | 11100 |

838

| | |
|---|-------|
| TCTAGCTTCC CCATTCTATG GAATCTTGCA TTATCCATAA TAATAACCGA TGGTGTGGTT | 11160 |
| AATGTTGGTA AGAGAAACTT CTGAAACCAA GCTTCAAAAA AGTCGCTCGT CATCGTCTCT | 11220 |
| TCGTAAGTCA TTGGAGCGAT TAACTCACCA TTTGTTAGAC CTGCAACCAA AGAAATCCTC | 11280 |
| TGATATCTTC TTCCAGATAC TTT | 11303 |

(2) INFORMATION FOR SEQ ID NO: 116:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3112 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 116:

| | |
|--|------|
| CCTTAGATTT CCACTTGCCA GAGGAATTGA TTGCCCCAAC GCCCCTTGAA AAACGTGATG | 60 |
| CCTCCAAACT CCTCATCGTC AACCGTGAGA CAGGAGAAAT GCAAGATAAA CATTTCCACT | 120 |
| CTATTATTGA TATGCTGGAA CCTGGTGATG CCCTTGTCAT GAACGACACC CGAGTTCTCC | 180 |
| CTGCCCCGCT CTATGGTCAA AAAGTGAGA CAGGAGGTCA TGTGGAAGTT CTCCTCCTTA | 240 |
| AGAACACTAG TGGAGACGAG TGGGAAGTTC TGGCTAAACC TGCCAAACGC CTCGAAGTCG | 300 |
| GTACTCGTAT CAGCTTTGGT GATGGCCGCC TCAGCGCTGT CGTTACAGAA GAATTGACCC | 360 |
| ACGGGGGACG CATGTGCCG TTTGAATACC AAGGAATTTT CCTAGAAGTC TTGGAAAGTC | 420 |
| TGGGAGAAAT GCCTCTGCCA CCTTATATCC ACGAAAAATT AGATGACCGT GAACGTTATC | 480 |
| AAACCGTCTA CGCCAAGGAA AGTGGCTCTG CTGCAGCACC GACTGCTGGT CTTCACTTCA | 540 |
| CCAAAGAACT GCTGGCAGAA ATCCAAGCTA AGGGTGTTC TCTAGTCTAT CTGACTCTCC | 600 |
| ATGTCGGACT CGGAACCTTT AGACCTGTTT CTGTGGATAA TCTGGACGAA CACGAAATGC | 660 |
| ACTCAGAGTT CTATCAACTT TCTGAGGAAG CTGCTGCCAC CCTTCGCTCT GTCAAAAAAA | 720 |
| ATGGTGGTCG TGTCATCGCT GTCGGAACCA CTTCTATCCG CACCTTGGA ACTATTGGTT | 780 |
| CCAAGTTTGA TGGGCAAATC CAAGCAGATT CTGGTTGGAC CAATATCTTT ATCAAACCTG | 840 |
| GGTATGAGTG GAAGGTCGTG GATGCCCTTCT CAACCAACTT CCACCTGCCA AAATCAACTC | 900 |
| TGGTCATGTT GGTTCCTGCC TTTGCAGGCC GTGAATTAGT CTTAGATGCC TACCACCATT | 960 |
| CCATCCAAGA ACACTACCGC TTCTTCAGTT TTGGTGACGC CATGTTTATT TATTGAGAAA | 1020 |
| GAATTTCTCT AAATCTTCTA ATACCAATAA ATCGCTAAGA TATTATTTCA AAGAACATCT | 1080 |
| ACAATTGAAA CTCTAGCTAG CTGTAGAAGA GGCCTAGTAC ATTGAAATTA AAATGCTTCC | 1140 |
| CCCTAGCTTC GAAAATATG CCATAGATTG CGTTGACTCT CCAAATTGAT TCATCTATAT | 1200 |

839

| | |
|--|------|
| TTTATTTTCAG CTCCTATAC TTTCTTCGCT GTTTGTAAAT CAAAATGCAA GACACATGAG | 1260 |
| TAGCACCATA TTTGTTACTC TTATCTGTCC TCTCAAGAGA CTATTATGAG TTATTTTCAGA | 1320 |
| ATCATTCACT ACTTTGACCC TGACTCTCCT TAGTCTCAAA ATCAAAGACT TATACTCTTC | 1380 |
| AAAAATCTCT TCAAACCGCG TCAACGTCAC CTTGGATTAT ATATGTGatC TGactTCGTC | 1440 |
| AGTTCTATCT ACAACCTCAA AGCAGTACTT TGAGCAACCT GCGACTAGTT TTCTAGTTTG | 1500 |
| CTCTTTGATT TTCATTGAGT ATTAAACAAA AAGTGAACAA ATCTGAATTC TAATGTACAG | 1560 |
| AAGACTAGGC TTGTTCACTT TTTTATAGTC GCTATAAGAT GACCTTATCT ATAGCTTTTT | 1620 |
| ATATATAATT ATATATTCAG ACATACTATT ATCAATTTTG TCGCAGGGAG GAATCTGTTA | 1680 |
| ACGCACCCAT TCACCATTAT CATTGACTCT ATAGCCATCT ATACTTGTAT TGACCGCTAA | 1740 |
| CTCACCCGAT GTATTTACAT AATACCATT ACCACCAACT TGGAAACCATT GATTGACTTT | 1800 |
| CATAGAACCG TTGCTGTTGA GGTAGTACCA TGAACATTA ACTTGTACCC AACCTGTTGC | 1860 |
| CATGGAACCA TCAGTATTAT AAAAATACCA CATACCATT TCTTGTTTCC AGTCTGTTGT | 1920 |
| TGGAGCAACT GCTTTAGCTG GTTCTACTGC TACATCTGTT CCTTGGTTAG ATGTAACAGA | 1980 |
| TACAGGATAC GAAGGAATAG ATGATTGCTC AGGAACAACA ACTTTTTCAG GTTCTCTCGT | 2040 |
| CCCTCTCCTT ATACGTCTTT TTACCATCTC TTTAGTAATT TGACGAGAAG TAGTTTCTTC | 2100 |
| AATTGTTCCA TCACGTTTAT CTACAGTATA GATTGTAGTA AGAGTAATTT ACCAATTTCT | 2160 |
| CCTACTTCTT CTACTTCTTG ACTTTTATCA AGAGTTGGGC CATCGAGATA TTCTGTTTCG | 2220 |
| ATTGGAATTT CTGGACAAG AACTTGGGGC TTGGTTCTTT TTTTAACAAC TCTTGTTTGA | 2280 |
| GAGTCTTTTT TTTGACTTAA AGTACTCTCA GTTACTTGTC CACTCTTCC ATCTACATTA | 2340 |
| TAAGTTATCG TTGTAACGT TTTCCCATTC TTTCCTAGAG TAATCTCTTG CTCCTGTCCT | 2400 |
| GCAGAAAGGT CATTGTCTGC TTCATATTTA GTAGCAAATG GAACAAGAAC TTCTTCAACC | 2460 |
| TTGCTTTTAG CTGGAACCTT GATAACTGTA TCCGTGGCTT CTTTCTATC AACAGTAACC | 2520 |
| TGTTCCGTAA CATAACCAGT CTCTGGATTA ACATCGTAGG TCCTTGTCGT AGTTACATAG | 2580 |
| CCATCCTCTC CATCAATTGT AACAGGATTT TCACTACGGT CTTTGTGTTT ATCTTTTCA | 2640 |
| TAACGAATTC GCGTACTTGA AATTTTCTTG GTTACTACCT TAGGTTTAGT CGCTACTTTT | 2700 |
| ACAATAATAT CCCCATTTGTC AGCGTCATCA TACTCTATTC CCTCTTCTTT ATCTCTAGTA | 2760 |
| TCATCTCTGA CATATTGAAT CCCATCAGCA GCATGAACAA AACTTGTATT CAGATTCCCTC | 2820 |
| CTAAAAATAA AGTTAGCCCG ATTACCGCAG AACC AAAAAT CTTTCCGAGT TTACGTATTG | 2880 |
| CATAGCGCTT ATTAGTATTA GATTTTGCCA TTACATCCTA CTCTAGTAT AGCATCTTTT | 2940 |

840

| | |
|--|------|
| CTATCAAACG TTAAACAATA TACGTTATAT ATAAAAATAGA CTTAGAATGA TATATTGATT | 3000 |
| ATTGAACTAA CACTTTAACT ATATCGTAAT CAATCTCATA TATAAAGGAT TGCAGACATC | 3060 |
| TTATCTAAAT ACATGCGAAT ATATTTAGAT ACAAACATTC CAACTTGATA AT | 3112 |

(2) INFORMATION FOR SEQ ID NO: 117:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 4327 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 117:

| | |
|---|------|
| CCCCAAAATC TCTTCAAACC ACGTCAGCTT CGCCTTGCCG TAGTATGGTT ACTGACTTCG | 60 |
| TCAGTTCTAT CCACAACCTC AAAACAGTGT TTTGAGCATC ATGCgGCTAG CTTCTTAGTT | 120 |
| TGCTCTTTGA TTTTCATTGA GTATAAAAC AGATGAGTTT CTGTTTCTT TTTATGGACT | 180 |
| ATAAATGTTT AGCTGAAACT ACTTTCAAGG ACATTATTAT ATAAAAGAAT TTTTGAAAC | 240 |
| TAAAATCTAC TATATTACAC TATATTGAAA GCGTTTAAA AATGAGGTAT AATAAATTTA | 300 |
| CTAACGCTTA TAAAAAGTGA TAGAATCTAT TTTTATGTAT ATTTAAAGAT AGATTGCTGT | 360 |
| AAAAATAGTA GTAGCTATGC GAAATAACAG ATAGAGAGAA GGGATTGAAG CTTAGAAAAG | 420 |
| GGGAATAATA TGATATTTAA GGCATTCAAG ACAAAAAAGC AGAGAAAAAG ACAAGTTGAA | 480 |
| CTACTTTTGA CAGTTTTTTT CGACAGTTTT CTGATTGATT TATTTCTTCA CTTATTTGGG | 540 |
| ATTGTCCCTT TTAAGCTGGA TAAGATTCTG ATTGTGAGCT TGATTATATT TCCCATTATT | 600 |
| TCTACAAGTA TTTATGCTTA TGAAAAGCTA TTTGAAAAAG TGTTCGATAA GGATTGAGCA | 660 |
| GGAAGTATGG TGTAATAGC ATAGGCTGAT GTCCATCATT TGCTTATAAA GAGATATTTT | 720 |
| AGTTTAATTG CAGCGGTGTC CTGGTAGATA AACTAGATTG GCAGGAGTCT GATTGGAGAA | 780 |
| AGGAGAGGGG AAAATTGGCA CCAATTTGAG ATAGTTTGTT TAGTTCATTT TTGTCAATT | 840 |
| AATGAACTGT AGTAAAAGAA AGTTAATAAA AGACAAACTA AGTGCAATTT CTGGAGTAAA | 900 |
| TGTCTTATTT CAGAAATCGG GATATAGATA TAGAGAGGAT CAGTATGAAT CGGAGTGTTT | 960 |
| AAGAACGTAA GTGTCGTTAT AGCATTAGGA AACTATCGGT AGGAGCGGTT TCTATGATTG | 1020 |
| TAGGAGCAGT GGTATTTGGA ACGTCTCCTG TTTTAGCTCA AGAAGGGGCA AGTGAGCAAC | 1080 |
| CTCTGGCAAA TGAACTCAA CTTTCGGGGG AGAGCTCAAC CCTAACTGAT ACAGAAAAGA | 1140 |
| GCCAGCCTTC TTCAGAGACT GAACTTTCTG GCAATAAGCA AGAACAAGAA AGGAAAGATA | 1200 |
| AGCAAGAAGA AAAAATTCCA AGAGATTACT ATGCACGAGA TTTGGAAAAT GTCGAAACAG | 1260 |

841

| | |
|--|------|
| TGATAGAAAA AGAAGATGTT GAAACCAATG CTTCAAATGG TCAGAGAGTT GATTTATCAA | 1320 |
| GTGAACTAGA TAAACTAAAG AAACCTGAAA ACGCAACAGT TCACATGGAG TTTAAGCCAG | 1380 |
| ATGCCAAGGC CCCAGCATTC TATAATCTCT TTTCTGTGTC AAGTGCTACT AAAAAAGATC | 1440 |
| AGTACTTCAC TATGGCAGTT TACAATAATA CTGCTACTCT AGAGGGGCGT GGTTCGGATG | 1500 |
| GGAAACAGTT TTACAATAAT TACAACGATG CACCCTTAAA AGTTAAACCA GGTCAGTGGA | 1560 |
| ATTCTGTGAC TTTACAGTT GAAAAACCGA CAGCAGAACT ACCTAAAGGC CGAGTGCGCC | 1620 |
| TCTACGTAAA CGGGGTATTA TCTCGAACAA GTCTGAGATC TGGCAATTTT ATTAAGATA | 1680 |
| TGCCAGATGT AACGCATGTG CAAATCGGAG CAACCAAGCG TGCCAACAAT ACGGTTTGGG | 1740 |
| GGTCAAATCT ACAGATTCGG AATCTCACTG TGTATAATCG TGCTTTAACA CCAGAAGAGG | 1800 |
| TACAAAAACG TAGTCAACTT TTAAACGCT CAGATTTAGA AAAAAACTA CCTGAAGGAG | 1860 |
| CGGCTTTAAC AGAGAAAACG GACATATTCG AAAGCGGGCG TAACGGTAAC CCAAATAAAG | 1920 |
| ATGGAATCAA GAGTTATCGT ATTCCAGCAC TTCTCAAGAC AGATAAAGGA ACTTTGATCG | 1980 |
| CAGGTGCAGA TGAACGCCGT CTCCATTCTGA GTGACTGGGG TGATATCGGT ATGGTCATCA | 2040 |
| GACGTAGTGA AGATAATGGT AAAACTTGGG GTGACCGAGT AACCATTACC AACTTACGTG | 2100 |
| ACAATCCAAA AGCTTCTGAC CCATCGATCG GTTCACCACT GAATATCGAT ATGGTGTGG | 2160 |
| TTCAAGATCC TGAAACCAA CGAATCTTTT CTATCTATGA CATGTTCCCA GAAGGGAAGG | 2220 |
| GAATCTTTGG AATGTCTTCA CAAAAAGAAG AAGCCTACAA AAAAAATCGAT GAAAAACCT | 2280 |
| ATCAAATCCT CTACCGTGAA GGAGAAAAGG GAGCTTATAC CATTCGAGAA AATGGTACTG | 2340 |
| TCTATACACC AGATGGTAAG GCGACAGACT ATCGCGTTGT TGTAGATCCT GTTAAACCAG | 2400 |
| CCTATAGCGA CAAGGGTGAT CTATACAAGG GTGACCAATT ACTAGGAAAT ATCTACTTCA | 2460 |
| CAACAAACAA AACTTCTCCA TTTAGAATTG CCAAGGATAG CTATCTATGG ATGTCCTACA | 2520 |
| GTGATGACGA CGGGAAGACA TGGTCAGCTC CTCAAGATAT TACTCCGATG GTCAAAGCCG | 2580 |
| ATTGGATGAA ATTCTTGGGT GTAGGTCCTG GAACAGGAAT TGTACTTCGG AATGGGCCTC | 2640 |
| ACAAGGGACG GATTTTGATA CCGGTTTATA CGACTAATAA TGTATCTCAC TTAGATGGCT | 2700 |
| CGCAATCTTC TCGTGTATC TATTAGATG ATCATGGAAA AACTTGGCAT GCTGGAGAAG | 2760 |
| CGGTCAACGA TAACCGTCAG GTAGACGGTC AAAAGATCCA CTC'TTCTACG ATGAACAATA | 2820 |
| GACGTGCGCA AAATACAGAA TCAACGGTGG TACAACATAA CAATGGAGAT GTTAAACTCT | 2880 |
| TTATGCGTGG TTTGACTGGA GATCTTCAGG TTGCTACAAG TAAAGACGGA GGAGTGACTT | 2940 |
| GGGAGAAGGA TATCAAACGT TATCCACAGG TTAAAGATGT CTATGTTCAA ATGTCTGCTA | 3000 |

842

| | |
|--|------|
| TCCATACGAT GCACGAAGGA AAAGAATACA TCATCCTCAG TAATGCAGGT GGACCGAAAC | 3060 |
| GTGAAAATGG GATGGTCCAC TTGGCACGTG TCGAAGAAAA TGGTGAGTTG ACTTGGCTCA | 3120 |
| AACACAATCC AATTCAAAAA GGAGAGTTTG CCTATAATTC GCTCCAAGAA TTAGGAAATG | 3180 |
| GGGAGTATGG CATCTTGTAT GAACATACTG AAAAAGGACA AAATGCCTAT ACCCTATCAT | 3240 |
| TTAGAAAATT TAATTGGGAA TTTTGTAGCA AAAATCTGAT TTCTCCTACC GAAGCGAACT | 3300 |
| AGAGAGATGG GCAAAGGAGA GATGGGCAAA GGAGTTATTG GCTTGGAGTT CGACTCAGAA | 3360 |
| GTATTGGTCA ACAAGGCTCC AACCCTTCAA TTGGCAAATG GTAAACAGC GACTTTCCTA | 3420 |
| ACCCAGTATG ATAGCAAGAC CTTGTTGTTT GCAGTAGATA AGGAAGATAT CGGACAGGAA | 3480 |
| ATTATTGGTA TAGCTAAAGG AAGCATCGAA AGTATGCATA ATCTTCCTGT AAATCTAGCA | 3540 |
| GGTGCCAGAG TTCCTGGCGG AGTAAATGGT AGCAAAGCAG CGGTGCATGA AGTTCAGAA | 3600 |
| TTTACAGGGG GAGTTAATGG TACAGAGCCA GCTGTTTCATG AAATCGCAGA GTATAAGGGA | 3660 |
| TCTGATTGCG TTGTAACCTCT TACTACAAAA AAAGATTATA CTTACAAAGC TCCTCTTGCT | 3720 |
| CAGCAGGCAC TTCCTGAAAC AGGAAACAAG GAGAGTGACC TCCTAGCTTC ACTAGGACTA | 3780 |
| ACAGCTTTCT TCCTTGGTCT GTTTACGCTA GGGAAAAAGA GAGAACAATA AGAGAAGAAT | 3840 |
| TCTAAACATT TGATTTTGTA AAAATGGCTC TTTGTCAACT GTAGTGGGTT GAAGTCAGCT | 3900 |
| AAGCTCGAGA AAGGACAAAT TTTGTCCTTT CTTTTTTGAT ATTCTAGAGCG ATAAAAATCC | 3960 |
| GTTTTTTTGA GTTTTCAAAG TTCCGAAAAC CAAAGGCATT GCGCTTGATA AGTTTGATGA | 4020 |
| GATTATGGT CGCTTCCAAT TTGGCGTTAG AATAGTGTAG TTGAAGGGCG TTGACGATTT | 4080 |
| TCTCTTTGTC CTTTAGAAAG GTTTTAAAGA CAGTCTGAAA AAGAGGATGA ACCTGCTTTA | 4140 |
| GATTGTCCTC AATGAGTCCG AAAAATTTCT CCGGTTTCCT ATTCTGAAAG TGAAACAGCA | 4200 |
| AGAGTTGATA GAGCTGATAG TGATGTTTCA AGTCTTGTGA ATAGCTCAAA AGCTTGTTTA | 4260 |
| AAATCTCTTT ATTGGTTAAA TGCATACGAA AAGTAGGGCG ATAAAAATGT TTATCGCTGA | 4320 |
| GTTTACG | 4327 |

(2) INFORMATION FOR SEQ ID NO: 118:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3521 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 118:

| | |
|--|----|
| CTCTGGCCCT GCCACTCCAA CGTTTTGTCA GGGTGCTTTT TTCATAAAGG AGTTCATTATG | 60 |
|--|----|

843

| | |
|--|------|
| TTAGATATCA AACGTATTCG TACAGATTTT GAAGCTGTCG CAGAAAAATT AGCTACACGT | 120 |
| GGTGTAGATG CTGCTGTCTT GAATGAAATG AAAGAAATCG ATGCTAAACG TCGTAACATC | 180 |
| TTGGTCAAGG TTGAAACTCT CAAAGCAGAA CGTAACACAG TTTCTGCTGA GATTGCCCAA | 240 |
| GCTAAGCGCA ACAAGGAAAA TACAGATGAC AAGATTGCTG CCATGCAAAA TCTATCTGCT | 300 |
| GAGGTTAAAG CCTTGGATGC TGAATTGGCA GAAATCGATG CTAAATTGAC AGAATTTACA | 360 |
| ACGACTCTTC CAAATATCCC AGCTGACAGC GTTCCTGTTG GGGCTGACGA AGACGACAAT | 420 |
| GTGGAAGTTC GCCGTTGGGG TACTCCACGC GAGTTTGA CTGAACCTAA AGCTCACTGG | 480 |
| GATCTCGGTG AAGACCTTGG TATCCTTGAC TGGGAACGCG GTGGTAAGGT AACAGGCGCT | 540 |
| CGCTTCCTCT TCTATAAAGG CCTCGGTGCT CGTTTGGAAC GTGCTATCTA CAACTTTATG | 600 |
| TTGGATGAAC ATGGAAAAGA AGGCTATACT GAAGTCATCA CACCTTACAT AGTCAACCAT | 660 |
| GATTCTATGT TTGGTACTGG TCAGTATCCA AAATTTAAGG AAGATACTTT TGAAGTCAGC | 720 |
| GATACCAACT TTGTCTTGAT TCCAACGCT GAAGTTCCTC TGACAAACTA CTACCGTGAT | 780 |
| GAAATCTTAG ACGGCAAAGA TCTTCCAATC TACTTCACTG CCATGAGTCC GTCATTCCGT | 840 |
| TCTGAGGCTG GTTCTGCCGG TCGTGATACG CGTGGCTTGA TCCGTTTGCA CCAATTCCAC | 900 |
| AAGGTTGAAA TGGTCAAATT TGCCAAACCA GAAGAATCTT ACGAAGAATT GGAATAAATG | 960 |
| ACAGCCAACG CTGAAAACAT TCTTCAAAAA CTCAACCTTC CATACCGTGT CGTTGCTCTC | 1020 |
| TCTACTGGAG ATATGGGCTT CTCAGCTGCG AAGACTTACG ACTTGGAAGT GTGGATTCCA | 1080 |
| GCACAAAACA ATTACCGTGA AATCTCAAGC TGTTCAAACA CAGAAGATTT CCAAGCCCGT | 1140 |
| CGTGCCCCAA TCCGTTACCG TGATGAAGCA GATGGCAAGG TGAAACTCCT TCATACCTTG | 1200 |
| AACGGTTCTG GACTTGCACT TGGACGTACA GTGGCTGCAA TTCTTGAAAA TTACC AAAAT | 1260 |
| GAAGATGGTT CTGTGACCAT CCCAGAAGCA CTTCTGCCAT ACATGGGTGG AGCTGAAGTC | 1320 |
| ATCAAACCAT AAAAAATAAG GTTTAGCTAT TTCTAGCTAG ACCTTTTTTC GTAACCAAT | 1380 |
| CAGATAAGCA CCTAGTACAA AGAATAAAAT AGTTAGGCAT ATAATGGTTT CAGCCAATAC | 1440 |
| CAGGTAATCC AGAAATGGAA GTTTCAAAAT TCCCTGAGCC ATCTTGAGCG AGGTCGCTGT | 1500 |
| GATAATGGTT GGAAGGTGA GGGCTGAGAA GGCTGGTTGA AAACCTTGTT TTAAATGTT | 1560 |
| GGGCAGACGA GTTAAACAA AGAAAAAGAA GGATTGAGAA GCCAAATCA TGACAATCAA | 1620 |
| GACCCAAGTC GGCAGGCTGG TTCCTCCTAC TCGAACTAGA GAAGCCAAGA GTAGAGAGAA | 1680 |
| AGGAGCACAG TAGATTCCTT CTTGTCCAAG CAAGGCTAGT GGGAGTGGAT GTTCTTTTAA | 1740 |
| ATCGCTATAA ATAAGGGGAT AGAGATAGAA GGTCAAGAGA AAACCAAAAC TCAAGGTCGC | 1800 |

844

| | |
|---|------|
| ATAGGCAATT TCGATAATAC CTACCAGAGG ATAGGTCAAG GCAGCCACTG CTATCCCCAC | 1860 |
| ATAGAGAACC GTCCAGCTTG GAGTGGCATG AACCTTCCGC CCTGGACAAG CAAACTTGAT | 1920 |
| GGTAAAACCA GCAATCAAGG TCAAAATCCA GAGAAATGAA AACCACCAA TCCCTTGTGC | 1980 |
| TACCAAAGGA AGATAAGAGA ATACGCGAAA GACATAGGTC GATAAAATCA TCCCAGCCAT | 2040 |
| AGGAAAGGTT GCCATTCTTG ACAAAGAGG GGGCTTGGTC AATTCTTGCT TGGTTTCTTT | 2100 |
| CCAATTAAAG AGATGCAGAA TTAGAAAGTA AATCCATAAA ACCAAACCAA TCAGACTAAA | 2160 |
| AAGATGGGAT AGAACCGGCA ACGTATCTAA AATAAGATTT CCAGCTCCTG CCAAACCTAG | 2220 |
| CAACAACCTT GAAATACTA AGGGGAGTTT TTTTCATCCTA ACCTCCAATA ATCATGTTAG | 2280 |
| TTTCAGTATA ACATAAAAGC GCTTAAATGA GGATTTAAAA AAACGAGTCC GCTTATTTCA | 2340 |
| GACTTCATTT TACTCAGATA TGAATTAGGC ATAAGGTTGC AATTCTGGAT TAATTGGTGT | 2400 |
| ATTAGCTAAG TTGTTGGCAT AGTTACAGAG GATTGCTAGG CTGACACCAA AAACCACATC | 2460 |
| CAAGGCATTT TGTGAGTGT AGCCAGCTTC TAAAACTCA GACAAGGCTT CATCTCCTAC | 2520 |
| ACGACCCTTG GTATTGATAA CTGCCAAGGT AAACCTAGCT AGGGTATCCA ATTTAGGATC | 2580 |
| TGTTTCAATT GGAGTACGAT TGCGAAGAGC TTGAATCAAG TCATCATTC A TCTGGATTG | 2640 |
| TTTGATGGAA AAGGCTGTGT GACCTGCGAC ACAGAAGGCA CAACCATTGG TCACGGCTGC | 2700 |
| CGTGATTGTC ACCACTTCAC GCTCAACGGG TGTGAGGCTG TTGCGACGGT GGATAGATGA | 2760 |
| GACAATTGCG TAGGCTTCTA AAACAGTCGG GGCATTGGCC AAGAGACCGA TTAGGTTGGG | 2820 |
| AATATAGCCA TTGTTGTCTT TTTCTACTGT TTCAAGAATT TCTTTCACTT CTGCTGGTGC | 2880 |
| TGACTCTACT GTATGGATAG TAAATGTTGT CATAAGATAC CTCTTTTCTT ATTATTGACA | 2940 |
| CTAATATTAT TGGAAAATCT TATAAAATCC TGATTCCTAA GTTTATCTAA GATAAAGCTT | 3000 |
| TATTCTCTCA TAAGATTTTC GTTGTATAT TAGTTTATCA CACTTCCAAT CACTTGTATA | 3060 |
| ATATATATTA TATATCAGGC TGATAAAAAT TATTTATAGG CAAAAAATC ACACGAGCTG | 3120 |
| TGTGATTCCA TTATTTGTCA AAATACTTTT TAGTTTCAGC AATAACGACT GCGACAAGA | 3180 |
| CCAAGAGGGC AATCAAGTTT GGCAGAGCCA TCAAGGCGTT AACGATATCT GCGATAATCC | 3240 |
| AGACCATATC CAACTCGATA AATCCTCCTA ACAAGACCAT GAGCACAAA ACCACACGGT | 3300 |
| AGAGCCAGAT AAAGCGAACC CCAAAGAGGA ACTCAAAACA GCGTTCTCCG TAATAGTTCC | 3360 |
| AACCTAGAAT CGTTGTAAAG GCAAAAAGTA CAAGGAAGAT GGTCAAGAGA GCAGGCCCAA | 3420 |
| AGTGTGAAAA GTTTGTTGAG AAAGCTGACT GAGTCAAGGC AACCCCATTC AAGTCACCGC | 3480 |
| TCCAAACTCC AGTTACCAAG ATGGTCAAAC CAGTTAGAGT A | 3521 |

(2) INFORMATION FOR SEQ ID NO: 119:

845

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1968 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 119:

| | |
|--|------|
| AACCTGGGCA AGCAAGCTAA AAGCAATGGG ACCTGGAATC CTAATGGCAA CTGCCGCTGT | 60 |
| TGGAGGTTCC CACATTGTAT CCTCAACTCA AGCTGGCGGT TCTTACGGTT GGTCTCTACT | 120 |
| TCTCTTGGTC ATCTTAGCCA ATGTCTTTAA ATATCCATTT TTCCGTTTGT GTGCTGAATA | 180 |
| CACAGCTGAT ACTGGAAAGA CTTTGGTTGA AGGTTATGCC GAAAAAGGAA AACTCTATCT | 240 |
| CTGGATTTTC TTTATCCTCA ATGTCTTTTC GGCTATGGTC AACACGGCTG GTGTTGCCAT | 300 |
| TCTGTGCTCA GCTATCATCG CCAGTGCCTT CCCAATGATT GGACTTAGCA TTA CTCACTG | 360 |
| GTCCCTCATT CTCGTTGCAA TCATTTGGGC TATGCTACTC TTTGGAGGCT ACAAACTTTT | 420 |
| AGACGGCATG GTCAAATGGA TTATGTCTGC CTTAACCATT GCGACTGTTC TTGCACTTAT | 480 |
| CATTGCGGCG GTCAAGCATC CAGAATACAG TTCTGATTTT GTCGAGAAGA CACCTTGGCA | 540 |
| AATGGCAGCT CTGCCCTTCA TCGTCTCCCT CCTAGGATGG ATGCCGGCTC CTATTGAAAT | 600 |
| TTCAAGCATC AATTCACCTT GGTCAGCTGA AAAGAGAAAG ACCGTCAACT TTAACACAGA | 660 |
| AGACGCTCTG TTTGACTTTA ACACTGGTTA TATTGGAACA GCTATCCTAG CCGTCTTCTT | 720 |
| TGTGGCACTG GGAGCACTGA TTCAGTATCC TACAGGGCAG GCGGTTGAAG CTGCTTCAGC | 780 |
| CAAATACATC TCTCAATTCG TGGGCATGTA TGCCTCTGTT CTTGGCGAAT GGTCCTGTTA | 840 |
| CTTGATTACC TTTATTGCCT TCCTCTGTAT CTTTGGAACA GTTATAACTG TTATCGATGG | 900 |
| CTATTCTCGC GTTAATCAGG AATCTCTCCG ACTGCTAATC AGTCAAAAAG AGGACAATCG | 960 |
| TAAATCTTTG AACATCTGGA TGACCATCAC TGCTATCATC GGTATCGTCA TTATCAAGTT | 1020 |
| CTTCGCTGGT CAGGTTTCAA CCATGCTCCG CTTTGCCATG ATTGGCTCTT TCCTGACAAC | 1080 |
| ACCTTTCTTT GCTCTTTTGA ATTACGCCTT GGTAACGCGT GAAAAACAAA ATCTTCCTTC | 1140 |
| TTGGCTCAAA CACCTTGCCA TTGCGGGATT GATTTTCCTC TTTGCTTCGC CATCTTCTTT | 1200 |
| ATCTACGCAC TCGCAATCGG AAAAGCAGGG TAAGGGACAA GCGCGAGATG AAGATAAGGT | 1260 |
| TTCAATTCAA GAGAAAATTC AGCAAAATAT TCTATGATAA AAAGCATAAG AACAAGGTTT | 1320 |
| TGAAGACCTG AACTTATGCT TTTTACGTT CTTAAAGACT GTTTATACTC AAAAAACAGT | 1380 |
| TGAACAACCT CAACCACCTC TTATAAGAAC TTTATACTAT TCGAGAATCT CTTCAAACCA | 1440 |

846

| | |
|---|------|
| CGTCAGCTCT ATCTGCAACC TCAAAGCTGT GCTTTGAGCA ACCTGCGACT AGCTTCCTAG | 1500 |
| TTTGCTCTTT GATTTTCATT GAGTATTAAT TCTCCTTTTC CAACTCATAA AAATCTGCGA | 1560 |
| TAATAGCTGC GACATGTTTG ATATCTTCCA GCATGCCTCG CATTTCAAAG TCAGCCAATA | 1620 |
| CAGGGAAGCC AAAGCGTTGA CTGTATTGCT TGGCTGTTAG GCAGTATTGG TTATTAAAGT | 1680 |
| TACGATTTCG TGACCCAACC ACACCAAAAC ACTTACTAGC ATTGTTACCA TAGGCAATAA | 1740 |
| AATCTCCAC CGGTGTCGTC AAAATCTCAA CATCTCCGTT ATCCACGCCA TTCCACCTT | 1800 |
| CGAGATAGGT CGGCAAAAAA GCGACATAGG GATGGTCCAT TTCATAGAAA TTTTGCCTT | 1860 |
| CCTTGACCAA ATCCTTGATA TGAATCTTTT GAACCTCAAT CCCTTTGTAC TGGGACAAGA | 1920 |
| GATAGTCTTT CAAGCGCGTC ACAAACCTTT CAGTGTGCC ACTCAAGG | 1968 |

(2) INFORMATION FOR SEQ ID NO: 120:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7172 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 120:

| | |
|--|-----|
| CCGCATTTTT TATCACTAGA CTCGAGACAT CTTTGTAGTG GCTCTTGCTC TCTGGTTTAA | 60 |
| TTTTCTTCCT TGCTCAAGGA CTCCTGCTAT TTCTCTGGT CGTCCGACTC AAACATCAAT | 120 |
| TCGCTGAGAT TTATCCTCAA ATCAATAAAA AGATTGCTT CTAATATTTA GGGGTCTCTA | 180 |
| CCATTGATTT TCTATTTTTT GTTCTCTTAG CCTTCATTAG TTCTCAGCGT TTTTCATCTC | 240 |
| TTATGCCAAT CATCACTGCT TGCCATTCTA CTTTTTATTA TATGACAGCT GACTACCTAA | 300 |
| GAGAAACTA TCCAGACTTT TACGACAAAC ACATCTCTTT ATGGGAGTGT CTCTAAAGAA | 360 |
| AAGGAGGTTT TAGCATGAAA AAAATCATCT TCATCAAAAC CATTCAACTC CTTGTCATTG | 420 |
| ATGGAATCAT GCTGGCATT TTAGACATTA AAAGGGGGCT TACTTGGGAC TGGATTTTGA | 480 |
| TTTATAGCGG TTGGCTCATT TTCTTTTCATC CTGTGCTATT GACCTATCTT TCAAACCAAC | 540 |
| TTTGTGACCA CTTTAGTTAA CTCTATTCCC AGATTAGACC GAGATTCTGG CGTTTGTCTT | 600 |
| TACAAATTCT CCTATGGGAT AGCCTGATGA TTCTCTCCTT GGTGTCTTTA AGTGATATTC | 660 |
| CACCTTTTCCT TCAGGGAAC CTCCTCATCC TAGGACATCT CATCCCTTCC TATCGCATCT | 720 |
| GCCAAAGCCT GAAAGAGAC TTCCCCCAAG CATATCAAGA ACCGATTTCT TTTTGGAGTA | 780 |
| TTTTATGATA GATGAGAAAG ACCAAGCCGA CTGGGCTTGG TCTTTCTTAT CTCTTTTATG | 840 |
| TATCTAGGAT AATGGTAACA GGTCATTAT TAACCAGCTC AACCTGCATA TCTGCTCCAA | 900 |

847

| | |
|--|------|
| AGATGCCTGT CTGAACGGGC ACTTCTTGCG CTAATTTTGG ATTGAAAGCA TCATAGAAGT | 960 |
| CTGATGCCAT ATCAGGTTTA GCTGCCCCCTG TAAAGgCTGG ACGATTGCCT CTCTTAGTAT | 1020 |
| CCGCAAAGAG GGTAACTGA GAAATAGAGA GGATTTCTCC TTCAATATCT TTGACAGACA | 1080 |
| GGTTCATCTT GCCTTCTGCG TCTGAAAAA TCCGCATATT GACCAGTTTT CTCACAGCAT | 1140 |
| AGTCCAAATC TTCCTCTTGG TCCTCTGGTC CAACACCAAC CAGCAATAAA AGTCCCTGAT | 1200 |
| TGATTTTTCCT CTGAATCTGG CCTTCTATAC TCACTTGGGC TTTTTTAACC CGTTGGATAA | 1260 |
| TGATTTTCAT AATAGCCTTT CTAGTAAGAG CTAGGACAAC TAGCCGTTGG TCCGTTTGAC | 1320 |
| AGAGTAAACT TCTGGCACAC TCTTAATTTT ATCGACAACC GTGGTCAGTG TAGAGAGGTT | 1380 |
| GGCAATACCG AAGgACACAT GGATATTAGC AACTTCATA TCCTTG GTTG | 1440 |
| GACCGTTGAA ATATTCTTGG TTGTATTGA AAGAACTGCG AGTACATCGT TCAACAGTCC | 1500 |
| TGTACGGTTG AGACCGTAGA TATCGATATG GGCCATATAC TCCTTATTTG AGCTAGGGTA | 1560 |
| CTGGTCTTCC CATTCCACAT CAAGGAGACG TTGCTCGTAG TTTTCTTGGG CACGCAGGTT | 1620 |
| CATACAGTCC ACACGGTGAA TAGCCACACC ACGACCCTTG GTAATGTAGC CAACAATATC | 1680 |
| GTCACCAGGC ACGGGGTAC AACACTTAGC AATCCGCACT AGGAGACCAG AAGCACCTTC | 1740 |
| AATAACCACT CCCCCCTCAT GCTTGACCTT GAGGGTTTCT TTATTTTCAA CCTTGACCTC | 1800 |
| GCCACCTTTG ACAAGCTCCT CTGCCTCAGC TTTGGCCTTG GCACGCTCTT CCTCACGGCG | 1860 |
| TTCCCTTTCA GTCAGACGGT TAAAGACGGT AATCGCACCG ATTTCCCCAA AACCAATGGC | 1920 |
| CGCAAAGAGG GAGTCTTCG TCTTGTAAC TGTCTTTTGC AGAACTTGAT CCATGTGGCG | 1980 |
| CTGTCCATA AATTTATTTG CCACATAGCC ATTTTCTTGG AACTGAGCCA TCAGCATCTC | 2040 |
| ACGACCCTTG TTGACAGACA ATTCCTTATC TTGGTTTTTA AAGAACTGGC GAATCTTATT | 2100 |
| GCGCGCCTTG CTAGTCTTGA CCATATTGAG CCAGTCACGG CTAGGTCCAA AGGAGTTTCGG | 2160 |
| GTTGGCGATA ATTTCAACCT GATCCCCTGT CTTTAACTTG GTTGTGAGTG GAACCATGCG | 2220 |
| GCCATTGACC TTGGCACCAG TTGCTTTTTC ACCGACCTTG GTATGGATTT CGTAGGCAAA | 2280 |
| ATCAATCGGT CCTGAATCTT TGGGAAGGGA ACGGACAGCT CCATCTGGGG TAAAAACGTA | 2340 |
| AATCTCCTCA GCCAAATAGT TTTCTTAAC AGAGTCCACA AATTCCTTAG CATCATCAGC | 2400 |
| CTGGTCTTGG AGCTCCATCA TCTCCTTGAT CCAGTTCATT CCAATAGCTG ATTCCTTGCT | 2460 |
| GTTAACTTGC CCCTTTATAC CTTTCTTATA AGCCCACTGA GCCGCAACCC CGTACTCAGC | 2520 |
| CACCTCGTGC ATTTCTTGG TTCGAATCTG GAATTCATC GGCCCTTTTG GTCCATAAAC | 2580 |
| AGTCGTATGG ATAGACTGAT AACCATTGGC CTGCGGTTG GCGATATAGT CTTTGAAGCG | 2640 |

848

| | | | | | | |
|-------------|-------------|------------|------------|------------|-------------|------|
| ACCTGGCATC | GGTTTCCAAA | ATTCATGCAC | GTAACCAAGC | ATGGCATAAA | CATCACTTTG | 2700 |
| GGTATCTAAA | ATACAACGAA | TAGCAATCAG | ATCATAGATT | TCCTCAAACC | GTTCCTCTCTT | 2760 |
| GTCCTGCATT | TTGCGGAAAA | TTGAGTAAAT | ATGCTTGGGA | CGACCATAAA | TCTTCCCTTT | 2820 |
| CAAGTGACGT | TCTGTCGTAT | ACTCCTCTAA | TTTGTGACT | ACCTCATCCA | CCAAGGCCTC | 2880 |
| ACGCTCCCCTG | CGCTTTTCCT | TCATCATATG | GGTAATCTTG | TAAAACTCCG | TTGGATTGAG | 2940 |
| ATAACGGAAA | GACAAGTCTT | CTAATTCCCA | TTTGACACTG | GAAATCCCCA | AACGATGGGC | 3000 |
| AAGCGGGGCA | TAGATTTCCTA | TGGTTTCTTT | GGAATACGC | TCCTGCTTGT | CTTTTCGAAG | 3060 |
| ATGTTTCAGG | GTCCGCATAT | TGTGCAAGCG | GTCAGACAGT | TTGACCAAAA | TAACGCGGAT | 3120 |
| GTCCTCAGAC | ATGGCCATGA | GCATCTTGCG | ATGATTTTCC | GCTAATTGCT | CCTCGATCGA | 3180 |
| TTTGTACTCG | ACCTTGCCAA | GCTTGGTAAC | TCCGTCAACA | ATCATCCGCA | CATCAGGACC | 3240 |
| AAACTCTCTT | TCCAAATCGT | CCAAAGTCGC | ATCTGTATCT | TCCACCACAT | CATGCAAGAA | 3300 |
| TCCACAAGCT | ACTGTTACAG | CATCCAGCTT | TAGCTTAGCT | AAAATACCTG | CCACTTGGAT | 3360 |
| AGGGTGAATG | ATATAAGGCT | CGCCTGATTT | GCGATATTGA | CCACTGTGGC | ATTCAACAGC | 3420 |
| ATAGACCAAG | GCCTTATGGA | CAAAATGAAC | ATCCTCTTCC | GTTAAATATT | CTTTGGTTAA | 3480 |
| AGCGACAAC | TCTTCGCCTG | TTAAATTAC | TTCTTTGCGC | ATCTCTACTC | TCCAATCTTT | 3540 |
| CCTACCATTT | TATCACTTTT | TTAAGAATAT | GAAAACCTAG | TTGGAACAGA | ATAAGAAAAA | 3600 |
| AATAATTCAA | AATTGCTTGA | TAATTCTGAA | TTATTGGTCC | GTAATATACT | ACGAAGTTAG | 3660 |
| ATTTTAACT | TAGGTGATAG | AAGGAGAGAT | AGAAGAACGG | AAACCATATT | GTAACCCAAA | 3720 |
| GACTTTCTGA | CTTCCCAAT | TCCATTGAAG | ATACGAAAGA | TAAACGGTGG | AACTCGTATC | 3780 |
| ACATACACTG | GTACCTTGAC | TGGATTTTGG | AATTAATACT | AAATGAAAAT | CAAAGAGCAA | 3840 |
| ACTAGGAAAC | TAGCCGCAGG | TTACTCAAAG | CACCGCTTTG | AGGTTGCAGA | TAAAGTTGAC | 3900 |
| GCGGTTTGAA | GAGATTTTGG | AAGAGTATAA | AAATCCTCAA | GATACTTTCT | TCTATCCTTT | 3960 |
| AGTTTATAAG | GAGAATACCT | ATGAAAAAAA | CTGCTATTTT | TATCTTTGCT | CTCCTAATGT | 4020 |
| TAGGAGTTTG | CTGCCTGTTT | CTATTCAGCC | AGCAAAGCTA | TAAAAAACAG | TCGTTCAATA | 4080 |
| CTATGCTAAC | GACCAGAACC | TGCCCAGTAG | GATAACTTAT | AGTGAATATA | GCGACAAATG | 4140 |
| AGAAGCCAAC | TACGGTAGCA | CTCTAAACAT | CACGTCTATC | AAACAAGCTA | ATGACGGAGT | 4200 |
| TTATGCAACC | TATGAAGGGC | AATTGACACC | TTTCCAATAT | TGATAAATTG | ATAACGAGCC | 4260 |
| TGTCTTCATC | TAGTCATGCT | GGTTTAAAG | TTCATTTTAA | ATCCTTACCT | ATTCTCCCTA | 4320 |
| ACTGTGCTAT | ACTTAATTTA | TACTCAATGA | AAATCAAAGA | GCAAACCTAG | AAGCTAGCCG | 4380 |
| CAGGCTGTTC | AAAGCACTGC | TTTGAGGTTG | CAGATAAAGT | TGACGCGGTT | TGAAGAGATT | 4440 |

849

| | |
|--|------|
| TTCGAAGAGT ATTAGTACAT TCTTTGAGAT TGGAGCTAGT ATGAAAATCC ATAAAACCGT | 4500 |
| GAATCCTGTT GCCTATGAAA ATACCTATTA TCTAGAAGGC GAAAAGCACC TCATCGTCGT | 4560 |
| CGATCCTGGT AGTCATTGGG AAGCCATTTCG TCAGACAATC GAGAAGATCA ACAAACCGAT | 4620 |
| CTGTGCTATT CTCTTGACCC ACGCCCATTA TGACCATATC ATGACTCTGG ACTTGGTTCG | 4680 |
| CGAGACGTTT GGCAATCCTC CTGTCTATAT CGCAGAGAGC GAAGCCAGCT GGCTCTACAC | 4740 |
| TCCTGTCGAT AATCTCTCCG GTCTCCCTCG CCACGATGAT ATGGCAGATG TGGTCACAAA | 4800 |
| ACCTGCAGAA CACACCTTTG TCTTTCACGA AGAATACCAA CTAGAGGAAT TTCGTTTAA | 4860 |
| GGTTCACCG ACCCCAGGGC ACTCTATCGG TGGTGTTCCT CTAGTCTTTC CTGATGCTCA | 4920 |
| TCTAGTCTTG ACGGGAGATG CTCTATTCCG CGAAACTATC GGACGGACCG ACCTTCCGAC | 4980 |
| TGGTAGCATG GAGCAACTCC TTCATAGTAT CCAGACCCAA CTCTTCACCC TACCAAACTA | 5040 |
| CGATGTCTAT CCAGGACATG GTCCAGCTAC TACTATCGCT CACGAAAAGG CCTTCAATCC | 5100 |
| CTTTTCTAG CAAGATGATG ACAATCGAAA TTAAAGTAA CTATCCAGCA AATCTTTCTA | 5160 |
| TTACAAAAGG CATCCTATCA AGGTTTTCAC ACATGATTGG ATGCCTTTTT TCTGATGACT | 5220 |
| AGATTTTTTG CATTACCAA TAATCACGCG CTCCTCTGGT GAACGCCACA TTCCGTCTCC | 5280 |
| TTCTTTGACA TCATAGGTTG TAAAGAAATC GTCGAAGTTT GGTACTTGCA CATTGACACG | 5340 |
| GAGTTTGGCT GGTGCGTGCA CATCGACGCT AGCCAAAAGT TTCATAAATT CTGGTCGACC | 5400 |
| TTTCATGCGC CAGATGCGAC CGAAGTTGTA GAAGAACTCT TCTGCTGAGA AGTCTGCTTC | 5460 |
| TCTCTAGCT GCTTCAAGCG CTGCTGCGAT TCCTCCCAAG TCAGCCACGT TTTCTGATAC | 5520 |
| AGTCAATTTA CCGTTAATGG TTGCTCCATA AGAATCCTGT CCATCAAATT GGTCAATGAC | 5580 |
| TTTTTGTGTT TTCTCCTTGA AGGCAGCATA GTCGCTCTCT GTCCACCAAT CCTTGAGGCT | 5640 |
| ACCATTTTCG TCAAAGGAAG CCCCGTTAGT ATCAAAGGCG TGGGAAATTT CATGGGCAAT | 5700 |
| CACTGCCCCA ATACCACCGT AGTTAGCAGA AGATGACTGA TGCAAGTCAT AGAAAGGCGC | 5760 |
| CTGTAAATG GCCGCTGGAA AGACAATCAG GTTCTTCTGA GGATTGTAGT AGGCATTGAC | 5820 |
| CATATGAGCA GGCATGCCCC ATTCTTATA ATCTACAGGC TGGTTCCACT TACTCCAACT | 5880 |
| GTGCTTGATT TCCACACGCG CAAAGGCTAG AGCATTCTCA AAAAGACTGG CAGTTTCATT | 5940 |
| CACTACCTTA TCCTTGTAAC GTGCAGGCAA TTCTTCTGGA TAGCCAATAT AAGGTTTGAT | 6000 |
| CACATTGAGC TTCACGATAG CCTGTTTACA GGTTCCTGGA GTGAGCCAGT CATTCTTAAG | 6060 |
| CAGACGCTCC TTATAAACAT CAATCATGGT TGCCACTTTT TTCTCCACAT CCGCCTTGGC | 6120 |
| TTCTGGAGAG AACTTCTCAC GGGCGTACCA AAGACCCAGG GCTTGCTTGA AAGGTTCTTG | 6180 |

850

| | |
|--|------|
| TGCTAGATGA TAAGCTGCTT TGACCTTATC TTTTGCCTCT GGAAGTCCAG AAAGGGCACG | 6240 |
| GCTGTAGGCA CCAGACAAAA CACGGATATC CTCTGTTAAA TAGCTGGTTG AAAGATTGAC | 6300 |
| AACACTCAAA ATCAAGGTTG CTTTAAGGAG AGACCAGGCT TCCTCACTGT AGAATTGCTC | 6360 |
| TGCTGCTTGC CAGAAACGTT CCTCGTCTAC AATAACCTTG TCTGGTAATT GCCCAATAAC | 6420 |
| TGCTTTGAAG AAGTCATCCA AAGGTAGGCG AGGCGCGAAT TTCTTGAAAT CTTGTAAGA | 6480 |
| ATATGGATGA TAGAGTTTAG CATATTCTGA ACTTTCTTCA TTAGAGAGCA CCACTGCCGC | 6540 |
| AACTCGGCGG TCCAATTCAA GTCTTTTTC TAGCAAGTCT TCAATTTCTT CATCAGAGAA | 6600 |
| ATCATAAGCC TTGAGGAGAT TTGCGCTGCT TTCTTTCCAA AGAGTCAAGA GCTCTTCGCG | 6660 |
| CTGAGGATGT TCTTCTGCAT AGTAGGTCGT ATCTGGCAAG ATTGTGCTTG GAGCGCTAGC | 6720 |
| CCATAGAACA TTGATTCTAG CATCCATAAA GTCTGGCGAT ACACCAAAAG GAAGGAAGTT | 6780 |
| TGGTTTTCCCT GCAAGCTCAA ACTCTGCTAG TTTAGCTGTA AAATCCGCAA AAGTCTCCAA | 6840 |
| TTCTTGGAAT TCTTTAAGGA GTGGTAAGAC AGGTGTGATA CCGTCAGCTT CTCTCTTGTC | 6900 |
| AAAATCACGA ACTAGGCGGT GGTATTTGAC AAAGTTTTC AAGATAGCAT CCTCAGGCAC | 6960 |
| TTCTTCACCT GCTAACCCT TGTCTGTGT CGCCAGCATC AGGTCTTCAA TTTCTTGCTC | 7020 |
| TAAATCAACA AAACCTCCTG TTTGAGACTT ATCTGCTGGG ATTTAGCTG TCTGTTGCCA | 7080 |
| TTCTCCATTG ATAGCATCAT AAAAAATC TGGATAACGT GTCATCTTGT TCTCGCTTTC | 7140 |
| ATTTGTATTT GCATTTATCT TAACAAAAAT CG | 7172 |

(2) INFORMATION FOR SEQ ID NO: 121:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 4518 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 121:

| | |
|---|-----|
| CGGGAAGTTA TGCGATCTAG ACTTCGTTCC TGTACAGCTA CTTTCTCAGG TGGTCTTGTT | 60 |
| GTTTGTATGA GTTTGTTTAG AGAGGATCTT TCTATGTCTT TCTTTCTTAT TTTTGTTTTA | 120 |
| TATGCTTTTC TGATTCTTTA TCTAATTAT GGTATTTTCA GACTAAAAAG GAAATACCGA | 180 |
| GTAGATGAAT AGCAAGGTTT TAGGTCTTCA GATTGATTTT TAGCACTCTT GATAAAAGAG | 240 |
| TGCTAATTTT TTGAGTTTTT GTCTTGACAT TCTCTTCTAA GGGTGTATAA TAGAATCATG | 300 |
| AGTTAGCACT TGGATGCATT GAGTGCTAAT TGATCAGACA GAGAGGAGTG ATGAGATGGT | 360 |
| TACAGAGCGT CAGCAGGATA TTTTAAATCT GATTATTGAC ATCTTTACCA AAACGCACGA | 420 |

851

| | |
|--|------|
| ACCTGTCGGA TCAAAAGCCT TGCAAGAGTC TATTAAGTCT AGCAGTGCAA CCATTGCTAA | 480 |
| TGACATGGCG GAACTAGAAA AACAAAGGTT GCTTGAGAAG GCTCATACTT CAAGTGGTCG | 540 |
| GATGCCAAGT GTTGCTGGTT TTCAGTACTA TGTGAAACAC TCACTGGATT TTGACCGGCT | 600 |
| GGCTGAAAAT GAGGTATATG AGATTGTCAA AGCCTTTGAT CAGGAATTCT TCAAATTGGA | 660 |
| GGATATTCTG CAAGAGGCTG CTAAGTACTT AACAGACCTG AGTGGCTGTA CGGTAGTGGC | 720 |
| ACTGGATGTT GAGCCGAGCA GGCAACGTTT GACAGCCTTT GATATCGTTG TTTTGGGGCA | 780 |
| ACATACAGCC TTGGCGGTAT TTACCCTAGA CGAGTCGCGA ACGGTTACTA GTCAGTTTCT | 840 |
| GATTCCAAGG AACTTCTTGC AGGAGGATTT GCTGAAACTG AAGAGCATCA TTCAGGAACG | 900 |
| TTTCCTCGGT CACACCGTTT TAGATATTCA CTACAAGATT CGGACGGAGA TTCCGCAGAT | 960 |
| TATCCAGCGT TACTTTACAA CAACGGATAA TGTCATCGAT CTCTTTGAAC ACATCTTTAA | 1020 |
| GGAAATGTTT AACGAAAACA TTGTGATGGC GGGCAAGGTC CATCTCTTGA ATTTTGCCAA | 1080 |
| TCTAGCAGCC TATCAGTTCT TTGACCAACC GCAAAAGGTG GCCTTGAGGA TTCGTGAGGG | 1140 |
| GTTGCGTGAG GATCAGATGC AAAATGTTTCG TGTTCGAGAC GGTCAAGAGT CCTGTTTAGC | 1200 |
| TGACCTAGCG GTAATCAGTA GTAAGTTCCT CATTCCTTAT CGGGGAGTTG GAATTCCTAGC | 1260 |
| CATTATCGGT CCAGTTAATC TGGATTACCA ACAGCTAATC AATCAAGTCA ATGTGGTCAA | 1320 |
| CCGTGTTTTG ACCATGAAGT TGACAGATTT TTACCGCTAC CTCAGCAGTA ATCATTACGA | 1380 |
| AGTACATTAA GATTGAAATC ATTAAGGAG GCGAACATGG CCCAAGATAT AAAAAATGAA | 1440 |
| GAAGTAGAAG AAGTTCAAGA AGAGGAAGTT GTGAAAACAG CTGAAGAAAC AACTCCTGAA | 1500 |
| AAGTCTGAGT TGGACTTGGC AAATGAACGT GCAGATGAGT TCGAAAACAA ATATCTTCGC | 1560 |
| GCTCATGCAG AAATGCAAAA TATCCAACGC CGTGCCAATG AAGAACGTCA AAAGTTGCAA | 1620 |
| CGTTATCGTA GCCAGGACTT GGCAAAAGCA ATCTTACCAT CTCTTGACAA CCTTGAGCGT | 1680 |
| GCACTTGACG TTGAAGGTTT GACAGATGAT GTGAAGAAGG GCTTGGGGAT GGTGCAAGAA | 1740 |
| AGCTTGATTC ACGCTTTGAA AGAAGAAGGA ATTGAAGAAA TCGCAGCAGA TGGCGAATTT | 1800 |
| GACCATAACT ACCATATGGC CATCCAACT CTCCCAGCAG ACGATGAACA CCCAGTAGAT | 1860 |
| ACCATCGCTC AAGTCTTTCA AAAAGGCTAC AAAGTCCATG ACCGCATCCT ACGCCCAGCA | 1920 |
| ATGGTAGTGG TGTATAACTA AGATATAAAG CCCGTAAAAA GCTCGCAGTA AAAATAGGAG | 1980 |
| ATTGACGAAG TGTTGATGA ACACAAGAAA ATCTATCTTT TTTACTCAGA GCTTAGGGCG | 2040 |
| TGTTGATTC GGCAATTCTG ACGGTAGCTA AAGCAACTCG TCAGAAAACG GCAATCGCTA | 2100 |
| TGGCGTTTGC CTAGCTTCCT TACTAACTCG TCGTCGAAAT AAAATCGATT TCGACTCCTC | 2160 |

852

| | | | | | | |
|-------------|------------|------------|------------|------------|------------|------|
| GTGTCGCAAT | TTACATAATA | GAAAACTTGT | CCGAAACGAC | AATAAACTAT | GAAGAAAGAT | 2220 |
| AAAATATGTT | TGGCTTTGTA | ATAGTGAGCG | AAGCGAACCA | AACACGATAC | TCTTCGCCGT | 2280 |
| GGCGCTATTT | GCGCAAATTT | TGAGACCTTA | GGCTCAAAGT | TTAGTCAAAG | AGATTGACGA | 2340 |
| AGTCAAGCTC | TGACGGCGTC | GCCACTGTCT | CCACTTAAGA | AGAGTATCAA | AAAGAAAAAT | 2400 |
| AGAAAAATTA | CTAACAAGGA | GAAAAACACA | TGTCTAAAAT | TATCGGTATT | GACTTAGGTA | 2460 |
| CAACAACTC | AGCAGTTGCA | GTTCTTGAAG | GAAGTGAAG | CAAAATCATC | GCAAAACCAG | 2520 |
| AAGGAAACCG | CACAACTCCA | TCTGTAGTCT | CATTCAAAAA | CGGAGAAATC | ATCGTTGGTG | 2580 |
| ATGCTGCAAA | ACGTCAAGCA | GTTACAAACC | CAGATACAGT | TATCTCTATC | AAATCTAAGA | 2640 |
| TGGGAAC TTC | TGAAAAAGTT | TCTGCAAATG | GAAAAGAATA | CACTCCACAA | GAAATCTCAG | 2700 |
| CTATGATCCT | TCAATACTTG | AAAGGCTACG | CTGAAGACTA | CCTTGGTGAG | AAAGTAACCA | 2760 |
| AAGCTGTTAT | CACAGTTCCG | GCTTACTTCA | ACGACGCTCA | ACGTCAAGCA | ACAAAAGACG | 2820 |
| CTGGTAAAT | TGCTGGTCTT | GAAGTAGAAC | GTATTGTTAA | CGAACCAACT | GCAGCAGCTC | 2880 |
| TTGCTTATGG | TTTGGACAAG | ACTGACAAAG | AAGAAAAAAT | CTTGGTATTT | GACCTTGGTG | 2940 |
| GTGGTACATT | CGACGTCTCT | ATCCTTGAAT | TGGGTGACGG | TGCTTTCGAC | GTATTGTCAA | 3000 |
| CTGCAGGGGA | CAACAAACTT | GGTGGTGACG | ACTTTGACCA | AAAAATCATT | GACCACTTGG | 3060 |
| TAGCAGAATT | CAAGAAAGAA | AACGGTATCG | ACTTGTCTAC | TGACAAGATG | GCAATGCAAC | 3120 |
| GTTTGAAAGA | TGCGGCTGAA | AAAGCGAAGA | AAGACCTTTC | TGGTGTAAT | TCAACACAAA | 3180 |
| TCAGCTTGCC | ATTTATCACT | GCAGGTGAGG | CTGGACCTCT | TCACTTGGA | ATGACTTTGA | 3240 |
| CTCGTGCGAA | ATTTGACGAT | TTGACTCGTG | ACCTTGTTGA | ACGTACAAAA | GTTCCAGTTC | 3300 |
| GTCAAGCCCT | TTCAGATGCA | GGTTTGAGCT | TGTCAGAAAT | CGACGAAGTT | ATCCTTGTGT | 3360 |
| GTGGTTCAAC | TCGTATCCCT | GCCGTTGTTG | AAGCTGTTAA | AGCTGAAACT | GGTAAAGAAC | 3420 |
| CAAAACAAATC | AGTAAACCCT | GATGAAGTAG | TTGCTATGGG | TGCGGCTATC | CAAGGTGGTG | 3480 |
| TGATTACTGG | TGATGTCAAG | GACGTTGTCC | TTCTTGATGT | AACGCCATTG | TCACTTGGTA | 3540 |
| TCGAAACAAT | GGGTGGAGTA | TTTACAAAAC | TTATCGATCG | CAACACTACA | ATCCCAACAT | 3600 |
| CTAAATCACA | AGTCTTCTCA | ACAGCAGCAG | ACAACCAACC | AGCCGTTGAT | ATCCACGTTT | 3660 |
| TTCAAGGTGA | ACGCCCAATG | GCAGCAGATA | ACAAGACTCT | TGGACGCTTC | CAATTGACTG | 3720 |
| ATATCCCAGC | TGCACCTCGT | GGAATTCCTC | AAATCGAAGT | AACATTTGAC | ATCGACAAGA | 3780 |
| ACGGTATCGT | GTCTGTTAAG | GCCAAAGACC | TTGGAAGTCA | AAAAGAACAA | ACTATTGTCA | 3840 |
| TCCAATCGAA | CTCAGGTTTG | ACTGACGAAG | AAATCGACCG | CATGATGAAA | GATGCAGAAG | 3900 |
| CAAACGCTGA | AGCCGATAAG | AAACGTAAAG | AAGAAGTAGA | CCTTCGTAAT | GAAGTAGACC | 3960 |

853

| | |
|---|------|
| AAGCAATCTT TGCGACTGAA AAGACAATCA AGGAAACTGA AGGTAAAGGC TTCGACGCAG | 4020 |
| AACGTGACGC TGCCCAAGCT GCCCTTGATG ACCTTAAGAA AGCTCAAGAA GACAACAAC | 4080 |
| TGGACGACAT GAAAACAAA CTTGAAGCAT TGAACGAAAA AGCTCAAGGA CTTGCTGTTA | 4140 |
| AACTCTACGA ACAAGCCGCA GCAGCGCAAC AAGCTCAAGA AGGAGCAGAA GGCGCACAAG | 4200 |
| CAACAGGGAA CGCAGGCGAT GACGTCGTAG ACGGAGAGTT TACGGAAGAA TAAGATGAGT | 4260 |
| GTATTGGATG AAGAGTATCT AAAAAATACA CGAAAAGTTT ATAATGATTT TTGTAATCAA | 4320 |
| GCTGATAACT ATAGAACATC AAAAGATTTT ATTGATAATA TTCCAATAGA ATATTTAGCT | 4380 |
| AGATATAGAG AATTATATTA GCTGAACATG ATAGTTGTAT CAAAAATGAT GAAGCGGTAA | 4440 |
| GGAATTTTGT TACCTCAGTA TTGTTGTCTG CATTTGTATC GGCGATGGTA CCGTATCTGA | 4500 |
| CGAACGTTCA GCTTATAT | 4518 |

(2) INFORMATION FOR SEQ ID NO: 122:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 8145 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 122:

| | |
|---|-----|
| TGCTATTTTC GATTCCCTTG GCGTTTTGA TTGCCTTTGC CTTGCAAGTC CATTGGAAGC | 60 |
| CCCTCCATTA TCTGATTAAC ATTTACATCT GGGTTATGCG AGGAACCCCC TTA CTCTTG | 120 |
| AACTGATTTT TATCTATTAT GTGCTCCCAA GTATTGGGAT TCGTTTAGAC CGCCTTCCTG | 180 |
| CAGCTATTAT TGCTTTGTT CTCAACTATG CAGCTTACTT TGCAGAAATT TTCCGTGGGG | 240 |
| GAATTGACAC TATTCCAAGA GGACAGTATG AGGCCGCCAA GGTCTTGAAG TTAGCCCTT | 300 |
| TTGACAGAGT GCGCTATATT ATCTTGCCCC AAGTGACCAA GATCGTTCTT CCTAGTGTCT | 360 |
| TTAATGAAGT TATGAGTTTG GTCAAGGATA CTTCCTTGGT CTATGCTCTC GGAATTCAG | 420 |
| ACCTTATCTT GGCTAGTCGA ACAGCTGCTA ACCGCGATGC TAGTCTAGTT CCTATGTTCT | 480 |
| TGGCAGGAGC CATTTATTTG ATTTTGATTG GGATTGTGAC AATTATTTCC AAAAAAGTTG | 540 |
| AGAAGAAGTA TAGTTATTAT AGATAGGAGG CTGCCATGTT AGAATTACGA AATATCAATA | 600 |
| AAGTCTTTGG AGACAAACAA ATCCTGTCTA ATTCAGTCT AAGTATTCCT GAAAAGCAAA | 660 |
| TCCTGGCTAT CGTTGGACCT TCTGGTGGAG GTAAGACAAC TCTTTTACGT ATGCTTGCAG | 720 |
| GTCTTGAAAC CATTGATTCA GGGCAAATCT TTTATAATGG ACAACCTTTA GAGCTGGATG | 780 |

854

| | |
|--|------|
| AATTGCAGAA GCGCAATCTA CTGGGATTGG TCTTCCAAGA TTTTCAACTA TTTCTCATC | 840 |
| TATCAGTTCT GGAAAATTTG ACTTTATCGC CTGTGAAGAC CATGGGAATG AAGCAGGAAG | 900 |
| AGGCTGAGAA GAAGGCGAGT GGACTCTTGG AACAGTTAGG ACTAGGAGGA CACGCAGAGG | 960 |
| CCTATCCTTT CTCACTATCT GGTGGGCAAA AGCAGCGGGT GGCTTTGGCG CGTGCTATGA | 1020 |
| TGATTGACCC AGAAATCATT GGCTACGATG AACCAACTTC TGCCCTGGAT CCAGAATTAC | 1080 |
| GTTTGGAAGT GGAGAAGCTA ATCTTGCAAA ATAGGGAAGT TGGGATGACC CAGATTGTGG | 1140 |
| TTACCCATGA TTTGCAGTTT GCTGAAAATA TCGCAGATGT ATTATTGAAA GTAGAACCTA | 1200 |
| AATAGGAGGA AAAATGGATG AAAAAATGGA TGCTTGTATT AGTCAGTCTG ATGACTGCTT | 1260 |
| TGTTCTTAGT AGCTTGTGGG AAAAAATCTA GCGAAACTAG TGGAGATAAT TGGTCAAAGT | 1320 |
| ACCAGTCTAA CAAGTCTATT ACTATTGGAT TTGATAGTAC TTTTGTTCCTA ATGGGATTTG | 1380 |
| CTCAGAAAGA TGGTTCTTAT GCAGGATTTG ATATTGATTT AGCTACAGCT GTTTTGAAGA | 1440 |
| AATACGGAAT CACGGTAAAT TGGCAACCGA TTGATTGGGA TTTGAAAGAA GCTGAATTGA | 1500 |
| CAAAAGGAAC GATTGATCTG ATTTGGAATG GCTATTCCGC TACAGACGAA CGCCGTGAAA | 1560 |
| AGGTGGCTTT CAGTAACCTA TATATGAAGA ATGAGCAGGT ATTGGTTACG AAGAAATCAT | 1620 |
| CTGGTATCAC GACTGCAAAG GATATGACTG GAAAGACATT AGGAGCTCAA GCTGGTTTAT | 1680 |
| CTGGTTATGC GGACTTTGAA GCAAATCCAG AAATTTTGAA GAATATTGTC GCTAATAAGG | 1740 |
| AAGCGAATCA ATACCAAACC TTTAATGAAG CCTTGATTGA TTTGAAAAAC GATCGAATTG | 1800 |
| ATGGTCTATT GATTGACCGT GTCTATGCAA ACTATTATTT AGAAGCAGAA GGTGTTTTAA | 1860 |
| ACGATTATAA TGTCTTTACA GTTGGACTAG AAACAGAAGC TTTTGCGGTT GGAGCCCGTA | 1920 |
| AGGAAGATAC AAAGTTGGTT AAGAAGATAA ATGAAGCTTT TTCTAGTCTT TACAAGGACG | 1980 |
| GCAAGTTCCA AGAAATCAGC CAAAAATGGT TTGGAGAAGA TGTAGCAACC AAAGAAGTAA | 2040 |
| AAGAAGGACA GTAAGATAAA ATAGTGGCTG AAAGTGCCTT TTGATTAGCA AAACGTAGTT | 2100 |
| TTTTTTGTAA TCTAGGAAAA CGATAATAGC GATTGAATAT GGATAATTGA ATATGGAATA | 2160 |
| GCCCACTGTG ATTTCTAAAA CATTGTTAAA AATTGATTGG ACTTCCAAAA TTAAATGTT | 2220 |
| CTGTAATGAA ATACTGATGT AACTGTTTTA GGAACAATAA AACGCATAAT ATCAAGGTTT | 2280 |
| TTGCACCTTA CATATGCGT TTTTGTGATT TTAAGACTTG TTAGCTGATT TTTTACAATC | 2340 |
| CTGCGAAATC TTTGATTCTT TGTGCTGACA TTGAAGAGTC GCAACGGACG TTGATTGTGC | 2400 |
| CATCTGTAAT ATGAACAAAA CCTGGTACAG TTGGGATFCC ATAGCGTGAG CGGAATGCTT | 2460 |
| GCAAATCATT GAGTTGGCTT GGTCTTTCAC TATTGATGAA GTAAATGTGA GCTTTGGTTT | 2520 |
| CAGCTACGAC ACCTGACAAT GTACCTGCAA ATTTACGGCA GTAAGGGCAA GTTTTGGCAG | 2580 |

855

| | |
|---|------|
| CGATAAAGAA GGTTCAGTT TCTTTTTTAT CAAGAGCTTC TTGCGCACGC ACAACTGTAG | 2640 |
| TGACTTCAAG GTCTTTGATG TTATCTAAAA ATTGTTCCAT GAGATTACCT CGCTTTCATT | 2700 |
| GATAAGTCTA GTATGCCATA AAGTTTCTAA AATTGCTTAG ATTTGATACG AAAAAAGATG | 2760 |
| AGGTTGGTTG GTCTCATCTT TTATAGGTCT TTATTTTACA AATGCATTGA TTTCTGCTTC | 2820 |
| GATGTTAGCA ATCTTAGCTT GTGATCTTC GTTGGTTTCC CCTACAACTG CAATGTAGAA | 2880 |
| CTTGATTTTT GGTCTGTAC CTGAAGGGCG AACGGCAATC CATGAACCGT CAGCAAGTGT | 2940 |
| GTATTTCAAC ACATCACTTG GAGGAGTTGT CAAGTTTGTA ACAGTACCGT CAGCAACAGT | 3000 |
| AGCAGTTTGT GCCTTGAAGT CTTCTACGAC AGTGATAGCT GTTGCCTTCC ATTCTGTTGG | 3060 |
| AGCATTGTTG CGGAATTTAG CCATAATCGC TTTGATTTGT TCAGCACCAT CGACACCTGA | 3120 |
| AAGAGTAACA GAGATTGTTT TTTCTGCGTA GTAGCCATAT TCTTTATAGA TTTCTTCGAT | 3180 |
| ACCGTCAGCA AGTGTCAAAC CACGAGAACG GTAGTAGGCA GCAAGTTCAG CAACTACAAG | 3240 |
| AACGGCTTGG ATGGCATCTT TATCACGTAC AAATGGTTTA ATCAAGTAAC CGAAGCTTTC | 3300 |
| TTCAAATCCC ATCATGTAAG TGTGGTTGTG TTTTCTCTCG AATCTTGGA TTTTTCAGC | 3360 |
| GATAAATTTG AAACCTGTCA AGACGTTGAA CATAGTTGCG CCGTAGCTTT CAGCAATCTT | 3420 |
| CGTTACCAAG TCAGTTGAAA CGATAGATTT GCAGAGAGCG GCATTTTCAG GAAGAGTTCC | 3480 |
| AGCGTTTTTG TGAGCTTCCA AGATGTATTT AGCCATGATA GCACCGATTT GGTACCTGA | 3540 |
| AAGGTTGAGG TAGCTACCAT CTTTTGAAG AACTTCAACA CCAACACGGT CAGCGTCTGG | 3600 |
| GTCAGTTGCG ACAAGAACAT CTGCACCAAC TTGACGACCA AGTTCTTCAG CAAGGGCAAA | 3660 |
| GGCTGCTTGG CTTTCTGGGT TTGGAGATGT TACAGTTGAA AAGTCTGGGT CAGCAGTTGC | 3720 |
| TTGCGCTTCA ACAACTTGAA CAGAGTCAAA TCCTGCTTGG GCAAGAGCAC GACGAGCCAA | 3780 |
| CATTTACCA GTACCATGAA GTGGTGTGTA GACAATCTTC ATGTCTTTAC CAAATCTTC | 3840 |
| AATCAAGGCT GGGTTGATGT TTATGTCCTT AACCTCTTTA AGGTATTCTA TGTCAACAGC | 3900 |
| TTGCGCGATA ACTTCAATCA AGCCAGAAGC TTTTTCAGTT TCCACATCAG CAACTTCAAC | 3960 |
| TGCAAATGGG TTTTCGATTG CACGGATATA AGTAGTCAAA GCGTCCGCAT CGTGTGAGG | 4020 |
| CATTGTGCCA CCGTCTTCAC CGTAAACCTT GTAACCGTTA AATGGAGCAG GGTGTGGCT | 4080 |
| GGCTGTGACC ATGATACCTG CGAAACAGTT GAGATGACGA ACTGCAAATG ATAGTTCTGG | 4140 |
| AGTCGGACGA AGGCTTTCAA ATACGTAAGA TTTGATGCCG TGTTTAGCAA GAACTGCCGC | 4200 |
| AGATTCAAAG GCAAACTCAG GTGAGAAGTG ACGGCTATCG TAGGCAATTG CTACACCGCG | 4260 |
| TTCTTTCTCG TTTCCACCTT TTGACTCAAT CAAACGAGCC AATCCTTCAG TAGCTTGGCG | 4320 |

856

| | |
|---|------|
| AACAACGTAG ATGTTGATAC GGTTTGTACC AGCACCAACC AAGCCACGCA TACCTGCAGT | 4380 |
| ACCAAATTCA AGATTTGTAT AGAAGGCATC TTCCTTAGTT TTTTCGTCCA TATTTTCCAA | 4440 |
| ATCTTGACGA AGGTAGTCAC GAAGCTCCAC AAAATCAACC CATTTCTGGT AATTTTCTTG | 4500 |
| GTAAGACATT CAAATCTCTCC TTTATTTTAA AAACATTAA TCAGTTAAT TATATCATTT | 4560 |
| TTTTTAGTTT TAGTAAAACC TTATCTGCTT CGAACATCTC TTCAAACCAG GTCAGATTGA | 4620 |
| ATTTTGGGGT TATATGATGT TGAGGCTAGG AAAAATTCAA TTTCAGTAAA AAAAGTAAGT | 4680 |
| CTTCTCATAA CAAACATGT ATATAGTTAC TTAGTTTAA ACAAGCATAT TATAATAAAG | 4740 |
| CTATGGCATA TAGTACTGAT TTAAACAGC GAGCATTAGA TTACATCAA GAGGGGCACA | 4800 |
| GCCATGTCGA GGCAGCCAAG TTTTGTGGTG TTGGCGTCAG AACTCTCTTC ACGTGGGAAA | 4860 |
| AGAAAGACGT GAACAAGAAC ACATAGAGAG GAAAAAGCGA GTCGTCAAAA ACCGAAAGAT | 4920 |
| TCCTTTAGAG GAATTGAAAG CCTTTGTAGA GGCTCATCCA GATGCTTTT TACGGGAAAT | 4980 |
| TGCGGCACAT TTTGATTGTG CTGTTCTTC AGTATGGCA GCTTTAAAGC AGATTAAGGT | 5040 |
| CACTTTAAAA AAAGATGACG AGCTTTAAGG AACAAGACCC AGAAAAGTAG CCTTATTCT | 5100 |
| TAAGAATTTT AATAGTTTAA AGCACCTAGC ACCTGTTTAT ATTGATGAAA CAGGAATCGA | 5160 |
| CCGCTATCTC TATCGTCCTT ATGCAGGGGC TCCTAGAGGG GAGAAAGTCT ATGAAAAGAT | 5220 |
| TAGCGGACGT CGTTTGTAGC GAACTTCAAT TGTTCAGGA CAAGTAGACG GAGAGTTTAT | 5280 |
| AGCTCCCATG ATTTACAAGA AAAGCATGAC AAGCGATTC TTTGTGGAGT GGTTCAAAAC | 5340 |
| GCAACTCCTA CCTGCTTTGA AGACACCTCA TGTATTGTC ATGGCAATG CTGGTTTCA | 5400 |
| TCCCAAGAAC ATTTTGGATG AACTCTGCAT CCAAGATAA CACTTTTCT TACCTCTACC | 5460 |
| ACCTTATTCA CCGGATTGA ATCCTATTGA GCAAGCTGG GCTATCTGA AAAAGAAAGT | 5520 |
| GACGGATGTA TTAAGGGAAG TTCCAACAT TTTTGAATGT TTGGAATGCT TTTTAAAAAC | 5580 |
| TAGATGACTA TAACGGTTCT AAAGGAACCT ATCGAGTAGT CATTAACCT AAGGATACTG | 5640 |
| CTGGTTAAGA GAAGACGGTA TACAATCAA CCATTCACCG TGAGCCGAA ATCGTTCAGA | 5700 |
| ATGAAGACTT GTATCAGAAT GAAGACTTGT ATAAGAAAGG TTTGAATGTT GAACTTGC | 5760 |
| ACCAACAAAT TAAGGGATTT TTTGAAGCAG AGTTTAAAA TCGTATTAAT GGAGTTCTTA | 5820 |
| ATACTAAAAA AAAAAATAGT ACATTAAATC GTGTAAATAA AAAAATAA CACCAGAGCA | 5880 |
| ACAAAACTC CATGATCAAT TTGAAGCAGA AGCAACGGAA GATGCTAAAA AACAAGCGA | 5940 |
| TATTGTGTTG AATGTTGACC AGGATTTTAT GAGCATATCT AAGTCTAATA AAAGTGGTTC | 6000 |
| AGACTGGAAG AAAACTTTCA CAGTGAGGAT AACCAATAGG CTAGCAAATG ACTTGAATAA | 6060 |
| TGTCTTGAAA CAGGTTGATA AAGATACTCC TAATACCCCA ACTTGGCTAA ACTCAGCTGC | 6120 |

857

| | |
|---|------|
| TTCTAAAGCT AAAGATGATG ACAGAGTATA TAACTACTG AAGACTCTTA TACCAGGAGA | 6180 |
| AAATTACCTA TCATGTTAAG GATAATCAGC TAGAAGTAGA AACAGATAAA TACACATATA | 6240 |
| CTGCCGCTAG AAATGGTAGT AAGGAAGTTG GTATTCAAGA GTCAGATATA GCAGCAACTC | 6300 |
| TAAGTGCCGA TGAATATAAT TCTAATCGCC AAACTTTGA GAGAGAATAC AAATACAAAA | 6360 |
| GCAATGCCC TTAATAATGG TTGGGCTAGA TCTGGTTCTG AAGAGTTCAA AAAGTTCTCC | 6420 |
| CACTTTGTAG GGGTAGACAA AGGGATTGTG CGAACGAATG TACTGACTGG TAAAAACTA | 6480 |
| TCTGATAAGA TTAGGAAAGA AGTGGGCTCT GGAGATAGCA AACTAGGAAA AGGCGGCTAT | 6540 |
| TTCTCTACTG GGGATGTTCT ATTAGGAAAA GATGTTGTTT CTTATACCGT ACAAGTATTT | 6600 |
| TCAGAGAATA ATGAAAGAGT AGGAGTAAAC ACTCAAAGTC ACCGTGTTCA GTATAATCTC | 6660 |
| CCAATTCTAG CTGACTTTTC AGTCATCCAA GATACTGTGG AACCATCAGC AACCGTTGTT | 6720 |
| GAAAAATCA TTCCAAACT AAATATTCCC GAAGAAGAGA AAGGGAAAT AACCGAAGAA | 6780 |
| ATCAAGAAAA AGAAAAAAC CTCAGAATTG GCAGAACTAA TCTCAGAAAA TGTGAAAGTT | 6840 |
| CGCTATGTTG ATGAACAAGG GCGTTTGCTA TCATTGAAAA ATGATACTGG AATTGGAGAA | 6900 |
| AAAGAAAGTG ACGGAACCTA CATTACCAAT AAAAAACAAC TGATTGGTAC CAGCTATAAT | 6960 |
| GTCACAGATA AAAAECTCAG TAGCATGACT ACTACTGACG GAAAATATTA TACTTTTAAA | 7020 |
| GAAGCAGATA CAAATTCTGC AAGTTTAACT GGGAATATTG TAAGCGAAGG TAGAACAGTG | 7080 |
| ACCTTAGTTT ATAGAGAAAG CGAAGCGCCA ACCACTGCTA CAGTAACAGC CAATTACTAT | 7140 |
| AAAGAAGGTA GGCAAGAGAA GTTGGTAGAG TCTGTTATAA AAGCTGATTT AGCGATAGGT | 7200 |
| TCTGAGTATA CCACAGAATC AAAAECTATT GAAGGGAAAA CAACAACTGA GGACAAAGAA | 7260 |
| GACCGAGTTA TCACAAGGAA AACACATAC ACCTTGGTAG CAACTCCTGA AAATGCGTAC | 7320 |
| CAGAAGACGG TGCAACAGTT GACTATTACT ACCGTGAGAA TGTGAGGAA ACAGTGGTTC | 7380 |
| CCAAAACAGC AACCTCTACT GAGACGAAGA CTATAACGCG TATCATTCAT TACGTTGATA | 7440 |
| AAGTTACGAA CCAAAATGTA AAAGAAGATG TTGTTCAACC TGTAACCTTA AGCCGTACAA | 7500 |
| AAACTGAGAA CAAGGTCACG GGAGTTGTAA CCTACGGTGA ATGGACAACA GGAACTGGG | 7560 |
| ACGAGGTTAT ATCTGGTAAG ATTGACAAGT ACAAAGATCC AGATATTCCA ACAGTTGAA | 7620 |
| CACAAGAAGT TACGTCAGAC TCTAGTGATA AAGAAATAAC GGTAAGGTAT GACCGTTTAT | 7680 |
| CAACACCAGA AAAACCAATC CCACAACCAA ATCCAGAGCA TCCAAGTGTT CCGACACCAA | 7740 |
| ACCCAGAACT ACCAAATCAA GAGACTCCAA CACCAGATAA ACCAACTCCA GAACCAGGTA | 7800 |
| CTCCAAAAAC TGAAACTCCA GTGAATCCAG ACCCAGAAGT TCCGACTTAT GAGACAGGTA | 7860 |

858

| | |
|--|------|
| AGAGAGAGGA ATTGCCAAAC ACAGGTACAG ³ AAGCTAATGC TACCTTGGCT AGTGCTGGTA | 7920 |
| TCATGACCTT GTTAGCTGGT CTAGGATTAG GATTTTTCAA GAAAAAGAA GATGAAAAAT | 7980 |
| AATAGATTTT AGAATCTAGG AACCAGGAAA AGCTCACAGA TGTGGGCTTT TTTCCTGGTT | 8040 |
| TTGAGAACGA GGTCTTTCGT AAAGAATAAA AACGCTTACA AGTCTGTTGA ACTGGGAAAC | 8100 |
| TATGAATCCT ATTTTTTTAA AAATATTTCC AGAAATCAGT TGC GG | 8145 |

(2) INFORMATION FOR SEQ ID NO: 123:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8697 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 123:

| | |
|--|------|
| CGGTACCGGG AACGATACTT AGTCTAATTT TGCACCTTTT CCATGTATGG TAAAGGTTTT | 60 |
| TCTTTTTTTA AAAAGGAAAA CGAGAAGAGG AGGTTCTTAT GAAAGCAAGC ATTGCCTTGC | 120 |
| AAGTTTACC CCTAGTACAG GGGATTGATC GGATAGCTGT TATTGATCAG GTCATTGCTT | 180 |
| ATCTGCAWAC TCAAGAAGTG ACGATGGTAG TGACACCATT TGAAACGGTC TTGGAAGGGG | 240 |
| AGTTTGATGA GCTTATGCGC ATTCTAAAAG AAGCGCTGGA AGTGGCAGGG CAGGAGGCAG | 300 |
| ACAATGTCTT TGCCAATGTC AAAATAAATG TAGGAGAGAT TTTAAGTATT GATGAGAAAC | 360 |
| TTGAGAAGTA TACTGAGACG ACACATTAGT CTATTGGGCT TTCTCGGAGT ATTGTCAATC | 420 |
| TGGCAGTTAG CAGGTTTCTT TAAACTTCTC CCCAAGTTTA TCCTGCCGAC ACCTCTTGAA | 480 |
| ATTCTCCAGC CCTTTGTTTCG TGACAGAGAA TTTCTCTGGC ACCATAGCTG GGCGACCTTG | 540 |
| AGAGTGGCTT TACTGGGGCT GATTTTGGA GTTTGTGATTG CCTGTCTTAT GGCTGTGCTC | 600 |
| ATGGATAGTT TGACTTGGCT CAATGACCTG ATTTACCCTA TGATGGTGGT CATTGAGACC | 660 |
| ATTCCGACCA TTGCCATAGC TCCTATCCTG GTCTGTGGC TAGGTTATGG GATTTTGCCC | 720 |
| AAGATTGTCT TGATTATCTT AACGACAACC TTTCCCATCA TCGTTAGTAT TTTGGACGGT | 780 |
| TTTAGGCATT GCGACAAGGA TATGCTGACC TTGTTTAGTC TGATGCGGGC CAAGCCTTGG | 840 |
| CAAAATCCTGT GGCATTTTAA AATCCCAGTT AGCCTGCCTT ACTTTTATGC AGGTCTGAGG | 900 |
| GTCAGTGTCT CCTACGCCTT TATCACAAC GTGGTATCTG AGTGGTTGGG AGGTTTGTAA | 960 |
| GGTCTGGTG TTTATATGAT TCAGTCTAAA AAAGTGTTC AGTATGATAC CATGTTTGCC | 1020 |
| ATTATTATTC TGGTGTGAT TATCAGTCTT TTGGGTATGA AGCTGGTCGA TATCAGTGAA | 1080 |
| AAATATGTGA TTAAATGGAA ACGTTCGTAG AATTAGAAATG TTTCTGAAAA AGAAAAGAGG | 1140 |

859

| | |
|---|------|
| AAATCAAAAT GAAGAAAACA TGGAAAGTGT TTTTAACGCT TGTAACAGCT CTTGTAGCTG | 1200 |
| TTGTGCTTGT GGCCTGTGGT CAAGGAACTG CTTCTAAAGA CAACAAAGAG GCAGAACTTA | 1260 |
| AGAAGGTTGA CTTTATCCTA GACTGGACAC CAAATACCAA CCACACAGGG CTTTATGTTG | 1320 |
| CCAAGGAAAA AGGTTATTTC AAAGAAGCTG GAGTGGATGT TGATTGAAA TTGCCACCAG | 1380 |
| AAGAAAGTTC TTCTGACTTG GTTATCAACG GAAAGGCACC ATTTGCAGTG TATTTCCAAG | 1440 |
| ACTACATGGC TAAGAAATTG GAAAAAGGAG CAGGAATCAC TGCCGTTGCA GCTATTGTTG | 1500 |
| AACACAATAC ATCAGGAATC ATCTCTCGTA AATCTGATAA TGTAAGCAGT CCAAAGACT | 1560 |
| TGGTTGGTAA GAAATATGGG ACATGGAATG ACCCAACTGA ACTTGCTATG TTGAAAACCT | 1620 |
| TGGTAGAATC TCAAGGTGGA GACTTTGAGA AGGTTGAAAA AGTACCAAAT AACGACTCAA | 1680 |
| ACTCAATCAC ACCGATTGCC AATGGCGTCT TTGATACTGC TTGGATTAC TACGGTTGGG | 1740 |
| ATGGTATCCT TGCTAAATCT CAAGGTGTAG ATGCTAACTT CATGTACTTG AAAGACTATG | 1800 |
| TCAAGGAGTT TGACTIONAT TCACCAGTTA TCATCGCAAA CAACGACTAT CTGAAAGATA | 1860 |
| ACAAAGAAGA AGCTCGCAAA GTCATCCAAG CCATCAAAAA AGGCTACCAA TATGCCATGG | 1920 |
| AACATCCAGA AGAAGCTGCA GATATTCTCA TCAAGAATGC ACCTGAACTC AAGGAAAAAC | 1980 |
| GTGACTTTGT CATCGAATCT CAAAAATACT TGTCAAAAGA ATACGCAAGC GACAAGGAAA | 2040 |
| AATGGGGTCA ATTTGACGCA GCTCGCTGGA ATGCTTTCTA CAAATGGGAT AAAGAAAATG | 2100 |
| GTATCCTTAA AGAAGACTTG ACAGACAAAG GCTTCACCAA CGAATTTGTG AAATAATGAC | 2160 |
| AGAAATTAGA CTAGAGCAGC TCAGTTATGC CTATGGTCAG GAGAGGATTT TAGAGGATAT | 2220 |
| CAACCTACAG GTGACTTCAG GCGAAGTGGT TTCCATCCTA GGCCCAAGTG GTGTTGGAAA | 2280 |
| GACCACCCTC TTTAATCTAA TCGCTGGGAT TTTAGAAGTT CAGTCAGGGA GAATTGTCCT | 2340 |
| TGATGGTGAA GAAAATCCCA AGGGGCGCGT GAGTTATATG TTGCAAAAGG ATCTGCTCTT | 2400 |
| GGAGCACAAG ACGGTGCTTG GAAATATCAT TCTGCCCTC TTGATTCAAA AGGTGGATAA | 2460 |
| GGCAGAAGCT ATTTCCCGAG CGGATAAAAT TCTTGCAGC TTCCAGCTGA CAGCTGTAAG | 2520 |
| AGACAAGTAT CCTCATGAAC TTAGCGGTGG GATGCGCCAG CGTGTAGCCT TACTCCGGAC | 2580 |
| CTACCTTTTT GGCACAAAGC TCTTCTCTT AGATGAGGCC TTTAGCGCCT TGATGAGAT | 2640 |
| GACAAAGATG GAACTCCACG CTGTTATCTT TGAGATTAC AAGCAGTTGC AGCTAACAAAC | 2700 |
| CCTGATCATC ACGCATAGTA TTGAGGAGGC CCTCAATCTC AGCGACCGTA TCTATATCTT | 2760 |
| GAAAAATCGC CCTGGGCAGA TTGTTTCAGA AATTAACTA GATTGGTCTG AAGATGAGGA | 2820 |
| CAAGGAAGTC CAAAAGATTG CCTACAAACG TCAAATTTTG GCGGAATTAG GCTTAGATAA | 2880 |

860

| | |
|--|------|
| GTAGAAAAAT AGGGAGTTGG TGAAGATTAT CCTTTACCAG CGCCCTTTT CTTTAAAAA | 2940 |
| TGAGAAAAAT TCGGTATAAT AGTCAAACAA GGTCAAGGTT TAAAGAGAGA GGTGGGTTTG | 3000 |
| TTATGAGATT TAAAAATACA TCGGATCATA TTGAGGCCTA CATCAAGGCG ATTTTAGATC | 3060 |
| AATCTGGTAT CGTGGAGTTG CAACGGAGTC AGTTGGCAGA TACCTTTCAG GTTGTTCCTA | 3120 |
| GTCAGATTAA CTACGTGATC AAGACACGCT TTACGAAAAG TAGAGGCTAC TTGGTTGAAA | 3180 |
| GTAAGCGTGG TGGCGGAGGC TACATTGTA TAGGACGGAT TGAGTTTCT AGTCATCATG | 3240 |
| AAATGCTCCG GGAGCTGCTT TACTCGATTG GTGAGCGAGT CAGTCAAGAA ATTTATGAGG | 3300 |
| ATATTCTCCA GCTTTTGGTT GAGCAGGAAT TGATGACCAA GCAGGAGATG AATTGCTAG | 3360 |
| AATCAGTAGC TTTGGATCGC GTTTTAGGAG AAGAAGCTCC AGTTGTTCGA GCAAACATGC | 3420 |
| TACGTCAGAT CATAAAGAG GTAGATAGAA AAGGGAAGTA AGATGAACTA TTCAAAGCA | 3480 |
| TTGAATGAAT GTATCGAAAAG TGCCTACATG GTTGCTGGAC ATTTTGGAGC TCGTTATCTA | 3540 |
| GAGTCGTGGC ACTTGTGAT TGCCATGTCT AATCACAGTT ATAGTGTAGC AGGGGCAACT | 3600 |
| TTAAATGATT ATCCGTATGA GATGGACCGT TTAGAAGAGG TGGCTTTGGA ACTGACTGAA | 3660 |
| ACGGACTATA GCCAGGATGA AACCTTTACG GAATTGCCGT TCTCCCGTCG TTTGCAAGTT | 3720 |
| CTTTTGTATG AAGCAGAGTA TGTAGCGTCA GTGTCCATG CTAAGGTA CTAGGACAGAG | 3780 |
| CACGTCCTCT ATGCGATTTT GCATGATAGC AATGCCTTGG CGACTCGTAT CTTGGAGAGG | 3840 |
| GCTGGTTTTT CTTATGAAGA CAAGAAAGAT CAGGTCAAGA TTGCTGCTCT TCGTCGAAAT | 3900 |
| TTAGAAGAAC GGGCAGGCTG GACTCGTGAA GATCTCAAGG CTTTACGCCA ACGCCATCGT | 3960 |
| ACAGTAGCTG ACAAGCAAAA TTCTATGGCC AATATGATGG GCATGCCGCA GACTCCTAGT | 4020 |
| GGTGGTCTCG AGGATTATAC GCATGATTTG ACAGAGCAAG CGCGTTCTGG CAAGTTAGAA | 4080 |
| CCAGTCATCG GTCGGGACAA GGAAATCTCA CGTATGATTC AAATCTTGAG CCGGAAGACT | 4140 |
| AAGAACAACC CTGTCTTGGT TGGGGATGCT GGTGTCGGGA AAACAGCTCT GCGCTTGGT | 4200 |
| CTTGCCAGC GTATTGCTAG TGGTGACGTG CCTGCGGAAA TGGCTAAGAT GCGCGTGTTA | 4260 |
| GAACCTGATT TGATGAATGT CGTTGCAGGG ACACGCTTCC GTGGTGACTT TGAAGAACGC | 4320 |
| ATGAATAATA TCATCAAGGA TATTGAAGAA GATGGCCAAG TCATCCTCTT TATCGATGAA | 4380 |
| CTCCACACCA TCATGGGTTC TGGTAGCGGG ATTGATTCTGA CTCTGGATGC GGCCAATATC | 4440 |
| TTGAAACCAG CCTTGGCGCG TGGAACTTTG AGAACGGTTG GTGCCACTAC TCAGGAAGAA | 4500 |
| TATCAAAAAC ATATCGAAAA AGATGCGGCA CTTTCTCGTC GTTTCGCTAA AGTGACGATT | 4560 |
| GAAGAACCAA GTGTGGCAGA TAGTATGACT ATTTTACAAG GTTTGAAGGC GACTTATGAG | 4620 |
| AAACATCACC GTGTACAAAT CACAGATGAA GCGGTTGAAA CAGCGTTAA GATGGCTCAT | 4680 |

861

| | |
|--|------|
| CGTTATTAA CCAGTCGTCA CTTGCCAGAC TCTGCTATCG ATCTCTTGA TGAGGCGGCA | 4740 |
| GCAACAGTGC AAAATAAGGC AAAGCATGTA AAAGCAGACG ATTCAGATTT GAGTCCAGCT | 4800 |
| GACAAGGCCC TGATGGATGG CAAGTGGAAG CAGGCAGCCC AGCTAATCGC AAAAGAAGAG | 4860 |
| GAAGTACCTG TCTACAAAGA CTTGGTGACA GAGTCTGATA TTTTGACCAC CTTGAGTCGC | 4920 |
| TTGTCAGGAA TCCCAGTTCA AAAACTGACT CAAACGGATG CTAAGAAGTA TTTAAATCTT | 4980 |
| GAAGCAGAAC TCCATAAACG GGTATCGGT CAAGATCAAG CTGTTTCAAG CATTAGCCGT | 5040 |
| GCCATTCGCC GCAACCAGTC AGGGATTGCG AGTCATAAGC GTCCGATTGG TTCCTTTATG | 5100 |
| TTCTAGGGC CTACAGGTGT CGGGAAGACT GAATTAGCCA AGGCTCTGGC AGAAGTTCTT | 5160 |
| TTTGACGACG AATCAGCCCT TATCCGCTTT GATATGAGTG AGTATATGGA GAAATTTGCA | 5220 |
| GCTAGTCGTC TCAACGGAGC TCCTCCAGGC TATGTAGGAT ATGAAGAAGG TGGGGAGTTG | 5280 |
| ACAGAGAAGG TTCGAATAA ACCCTATTCC GTTCTCTCTT TTGATGAGGT AGAGAAGGCC | 5340 |
| CACCCAGATA TCTPTAATGT TCTCTGCAG GTTCTGGATG ACGGTGCTTT GACAGATAGC | 5400 |
| AAGGGACGCA AGGTCGATTT TTCAAATACC ATTATCATTA TGACATCGAA TCTAGGTGCG | 5460 |
| ACTGCCCTTC GTGATGATAA GACTGTTGGT TTTGGGGCTA AGGATATTCG TTTTGACCAG | 5520 |
| GAAATATGG AAAACGCAT GTTTGAAGAA CTGAAAAAG CTTATAGACC GGAATTCATC | 5580 |
| AACCGTATTG ATGAGAAGGT GGTCTTCCAT AGCCTATCTA GTGATCATAT GCAGGAAGTG | 5640 |
| GTGAAGATTA TGGTCAAGCC TTTAGTGGA AGTTTGACTG AAAAAGGCAT TGACTTGAAG | 5700 |
| TTACAAGCTT CAGCTCTGAA ATTGTTAGCA AATCAAGGAT ATGACCCAGA GATGGGAGCT | 5760 |
| CGCCCACTTC GCAGAACCCT GCAAACAGAA GTGGAGGACA AGTTGGCAGA ACTTCTTCTC | 5820 |
| AAGGGAGATT TAGTGGCAGG CAGCACACTT AAGATTGGTG TCAAAGCAGG CCAGTTAAAA | 5880 |
| TTTGATATTG CATAAAGAA TAAAAGTATC AGCATCTGAC CATAAGTCAC AGTGGAGTGA | 5940 |
| AATTCAATGA AATCAAAGA GCAAACTAGG CAGCTAGCCG CAGGTTGCTC AAAACACTGG | 6000 |
| TTTGAGGTTG CAGATAGAGC TGACGTGGTT TGAAGAGATT TTCGAAGAGT ATGAAACTAA | 6060 |
| AACCTATAGC TTCTAAACGA TCCGTGGTTT TCATCATTTCA ACACAAAATT CATATGTTTA | 6120 |
| TTACCCCTCG TCGTATTTGT CTTAGAGCGT GTGTAGTAGA AAAAGAGCAG TCTTATCTGA | 6180 |
| AATTTTATTT CTTTCAAAAG AGACCTGTTT CTTTTTGTCA TGTCAAATCC GTTCTAGCTG | 6240 |
| GTATTTGAAA AATCAAATA ATATCAATG AAAATCAAAG AACAACTAG GAAGCTAGCC | 6300 |
| GCAGGTTGCT CAAAACACTG TTTTGAGGTT GTAGATAGAG CTGACGTGGT TTGAAGAGAT | 6360 |
| TTTCGAAGAG TATAAGCTGC AAGATGAATG ATTTTCTTGT ATTGACGTTG TTGTTGACAA | 6420 |

862

| | |
|---|------|
| AAAGTAGCGG ATAAATGAAA TCCATTCCAT TATCATAGAT GATAGGCTGG TAGGAAATTT | 6480 |
| TCAAATAGCA TACAGGAAAT AGATGTATGG AGTTCTGGTA GTAGAAAGGG AGAGAGATGA | 6540 |
| ACATTTTAGT TGCAGATGAC GAGGAAATGA TTAGAGAAGG AATTGCAGCA TTTCTGACAG | 6600 |
| AAGAGGGTTA TCATGTCATT ATGGCTAAGG ATGGACAAGA GGTCTTGGAA AAATTTCAAG | 6660 |
| ATCTCCCTAT CCATCTCATG GTACTGGATT TAATGATGCC TAGGAAGAGT GGTTTGAAG | 6720 |
| TGTTAAAGA AATCAATCAA AAGCACGATA TTCCTGTCAT CGTCTTGAGT GCTCTGGGAG | 6780 |
| ATGAACTAC TCAGTCACAG GTATTTGATC TCTATGCTGA TGATCATGTG ACAAACCTT | 6840 |
| TTTCTTTGGT ACTGCTTGTG AAGCGTATTA AGCGCTTAT CAGACGTTAC TACGTCATAG | 6900 |
| AGGATCTTTG GCGATATCAG GATGTAACAG TGGATTTTAC CTCTTACAAA GCACATTATA | 6960 |
| AAAATGAAGA AATTGATCTC AAACCAAGG AATTACTGGT ACTAAAGTGT TTGATTCAGC | 7020 |
| ATAAAAATCA AGTTTAAAGT AGAGAGCAGA TATTGGAAGA AATTTCAAAA GATGTAGCTG | 7080 |
| ATTTACCTTG TGATAGGGTC GTTGATGTCT ATATTCGTAC TCTTCGCAA AAATFAGCTT | 7140 |
| TAGATTGTAT CGTGACTGTG AAAAATGTTG GGTATAAGAT TAGCTTATGA TAAAAATCC | 7200 |
| TAAATTATTA ACCAAGTCTT TTTTAAGAAG TTTTGCAATT CTAGGTGGTG TTGGTCTAGT | 7260 |
| CATTCATATA GCTATTTAT TGAACCTTCC TTTTATTAT ATTCAACTGG AGGGGGAAAA | 7320 |
| GTTTAAAGAG AGCGCAAGAG TGTTTACGGA GTATTTAAAG ACTAAGACAT CTGATGAAAT | 7380 |
| TCCAAGCTTA CTCCAGTCTT ATTCAAAGTC CTTGACCATA TCTGCTCACC TTAAAGAGA | 7440 |
| TATTGTAGAT AAGCGGCTCC CTCTTGTCGA TGACTTGGAT ATTAAAGATG GAAAGCTATC | 7500 |
| AAATTATATC GTGATGTTAG ATATGTCTGT TAGTACAGCA GATGGTAAAC AGGTAACCGT | 7560 |
| GCAATTTGTT CACGGGGTGG ATGTCTACAA AGAAGCAAAG AATATTTTGC TTTTGTATCT | 7620 |
| CCCATATACA TTTTGGTTA CAATTGCTTT TTCCTTTGTT TTTCTTATT TTTATACTAA | 7680 |
| ACGCTTGCTC AATCCTCTTT TTTACATTTT AGAAGTGACT AGTAAATGC AAGATTTGGA | 7740 |
| TGACAATATT CGTTTGTATG AAAGTAGGAA AGATGAAGTT GGTGAAGTTG GAAAACAGAT | 7800 |
| TAATGGTATG TATGAGCACT TGTGAAGGT TATTTATGAG TTGGAAGTC GTAATGAGCA | 7860 |
| AATTGTAAAA TTGCAAAATC AAAAGGTTTC CTTTGTCCGC GGAGCATCAC ATGAGTTGAA | 7920 |
| AACCCCTTTA GCCAGTCTTA GAATTATCCT AGAGAATATG CAGCATAATA TTGGAGATTA | 7980 |
| CAAAGATCAT CCAAAATATA TTGCAAAGAG TATAAATAAG ATTGACCAGA TGAGCCACTT | 8040 |
| ATTAGAAGAA GTACTGGAGT CTTCTAAAT CCAAGAGTGG ACAGAGTGTG GTGAGACCTT | 8100 |
| GACTGTAAAG CCAGTTTATG TAGATATTTT ATCAGCTTAT CAAGAATTAG CTCATTCAAT | 8160 |
| AGGTGTTACA ATTGAAAATC AATTGACAGA TGCTACCAGG GTCGTCATGA GTCTTAGGGC | 8220 |

863

| | |
|---|------|
| ATTGGATAAG GTTTTGACAA ACCTGATTAG TAATGCAATT AAATATTCAG ATAAAAATGG | 8280 |
| GCGTGTAATC ATATCCGAGC AAGATGGCTA TCTCTCTATC AAAAATACAT GTGCGCCTCT | 8340 |
| AAGTGACCAA GAACTAGAAC ATTTATTTGA TATATTCTAT CATTCCTCAA TCGTGACAGA | 8400 |
| TAAGGATGAA AGTTCGGT TGGGTCTTTA CATGTGGAAT AATATTTTAG AAAGCTATCA | 8460 |
| AATGGATTAT AGTTTCTCC CTTATGAACA CGGTATGGAA TTAAAGATTA GCTGTAGAC | 8520 |
| AGATTAGTTT TTTATTAAAG TTCATATAGG GTTAACATAA GTGTGTTATT CTTGTGTAG | 8580 |
| ATAAAAGAAA GGATACTAAT ATGGTATTAG CGATTATTTT AGTAACATTC TTTATTCGAT | 8640 |
| TGATTTTTTTT AAAGCGTTCG ATAGAGAATG AGAAACGAAT CCTTAGCAAT GGCGGGG | 8697 |

(2) INFORMATION FOR SEQ ID NO: 124:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4317 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 124:

| | |
|--|-----|
| AACCATACAT ACGCAAGGC AAAGCTGACG CGGTTTGAAG AGATTTTCGA AGAGTATTAG | 60 |
| TTGCCTTTAA AGGCATCCAC CATCGTTTGA AATCTTCAT TTGAGAGAGT AATCCCTTTG | 120 |
| CCCATTTTAG TATGGTCTGG ACTCCAAGCA CGAATATCAA ACTTTGCAGG GGCACCATTA | 180 |
| AAGCTCACAC GGTAAATTC CTTGGTCCAA CCTTTTTCGT TTTCAGAAAG AGTCAACAAG | 240 |
| TGCTCTTCGA TTTCAAATGT AAATCTGCC ATTTTCTTCT CCTTTTTTAG TTTCATTAGT | 300 |
| TTATTCGTAA AATCTGTAG ATTTTAGGAA AATTTTATAT AATATTGATA TAAAGAAGG | 360 |
| GAGGCCAATA TGAGACATAA ATTCCAGCAA GTTCTAAATA AAATACATGA TTTTTTAAAT | 420 |
| GGATATGACC AACCTGACCA GACTGAAACC AACTCCCTTA CAGCCACTAT TGAAGAGGCT | 480 |
| ATCCAGAAAC AAACCGCTGT TCACCTTATC TTGTCTGAGA CAAGCTTTAC AGGTGACATC | 540 |
| ATCAAAATATG ATCAGCAAGG CCAGCAAATT ATCGTGAAAA ATTTTTCCAA AAATGTGAGC | 600 |
| CGGATTATCC GTATAAGCGA TATTCAACGC CTGCGATTG TCCCCTCAAC TGTCCTAAACA | 660 |
| GCCCCAAAAA ATAGATTTAA GAAAGAGTGA GATGTAGTTG CTTTATCCCA CTCTTTTTC | 720 |
| TTAGCGAATT TGTTCAAAAT GTAAATGAAC TGCGATATGA TCTCCATAAC CACTTCTTTC | 780 |
| CAAGTCACGT TGTAACGAT AGGAAATGTA GTGTTCTGCA ATGGTAATGT AACCTGCGCC | 840 |
| CAATAAACGA TGTTCACCA TAGATTGAAT CATACTGATA GTCGCACGTT CCACCTTGGC | 900 |

| | | | | | | |
|-------------|------------|------------|-------------|-------------|------------|------|
| 864 | | | | | | |
| TTCTTGTA | AACTA | CCTTCTTAGT | GACTTGAGCA | AGATTTTGAC | GCAAATCATC | 960 |
| TGTCAAACA | TAAACAGTTT | GGGCTGCCTT | CAAGATGGCT | TGGTAAATCT | TATCTGGATT | 1020 |
| AAATTCAGCA | ATTTGCCAT | TACGTTTGAT | TACTTGCCATA | GGTTTCTCCT | TTATTCTTTG | 1080 |
| TTTCTTTGA | TTTCTGCCAG | CATTTTTTCT | TCTTCTACTG | TCAGTTGATA | ATGTTCAAGT | 1140 |
| AAATCCGGTC | TGCGCTCGTA | GGTTTTCTTT | AAACTCTCGT | ACAATCGCCA | CTGACGAATC | 1200 |
| TTTTCATGGT | GGCCACTCAT | CAATACATCT | GGCACGACCA | TGCCTCGATA | ATCATAGGGA | 1260 |
| CGTGTGTACT | GAGGATATTC | TAAAGACCT | GAAGAAAAAC | TATCATCTTG | GTGGCTAGAC | 1320 |
| TCCTTGCCAA | TCACTTCTGG | AATCAGGCGA | ACTGTAGCAT | CAATCATGGT | CATAGCTGCC | 1380 |
| AATTCTCCAC | CAGTGAGGAC | ATAGTCACCT | AGGGAAATCT | CATCTGTTAC | CAAGGTCTTA | 1440 |
| ATGCGCTCAT | CATAACCCCT | ATAGTGCCCA | CAGATAAAGA | TTAGCTCTTC | CTCTTGAGCC | 1500 |
| AAATCTTCAG | CATAAGCCTG | ATCAAACGTC | TTCCAGCAG | GATCAAGGAG | AATAACGCGC | 1560 |
| GGATTTTTCT | TTTCAATAGC | ATCAAAGGAA | TCGAAAATAG | GTTGTGCTCT | GAGCAACATG | 1620 |
| CCCTGACCGC | CTCCGTAGGG | CTCATCATCT | ACATGACGGG | CCTTTTCAGC | ATTTTCTCGA | 1680 |
| AAATTATGAT | ACTGGATATC | CAAGAGCCCT | TTTTCTCGAG | CCTTTCCAAC | GATTGAGTGC | 1740 |
| TCCAGTGGAG | AAAACATCTC | TGGAAAGAGG | GTTAAAATAT | CAATCTTCAT | CGTCTAACCC | 1800 |
| TTCTAAGATT | TCCACATCGA | CCCGTTTACT | TGGAATATCA | ACATTGAGAA | CCACTGGTGG | 1860 |
| GATATAAGGT | AAAAGCAAAT | CACGTTTGCC | TTTTCGTTTG | ACCACCCAGA | CATCATTAGC | 1920 |
| ACCTGGTTGC | AGGATTTCTT | TGATGGTTCC | AACCAAGCTA | TCACCCCTCAT | AGACTTCCAA | 1980 |
| ACCGATAATC | TCGTGATAGT | AAAATTCACC | ATCGTCTAGG | TCATTCAAAT | CTTCTCAGC | 2040 |
| GACCTTGAGA | CTGTATCCCT | TGTACTTTTC | GATAGTATTG | ATATGGTACA | TATCTTTGAA | 2100 |
| TTTAATAATG | TCAAAGTTCT | TCTGTTTACG | GTGGCTAGCG | ATGGTCACTG | TTTGGACAAA | 2160 |
| CTGATCTTTT | TCATCAAACA | AAACCAGCTC | AGCTCCTTTT | TTAAACCGTT | CTTCTGCAAA | 2220 |
| ATCCGTCACA | GACAAGACTC | GCATCTCCCC | CTGTAATCCC | TGCGTATTAA | CGATTTTCCC | 2280 |
| AACATTAAAG | TAGTTCATCT | TGTCTCCTGT | AATCTCCTTT | TTCCATCTT | ATTCTAACAA | 2340 |
| TTCTCGAATA | ATAGCCGCAA | TTTTTTCCGA | TTCTGACCAT | TGTAAATAAT | GGTGATTCCC | 2400 |
| TCCTAAAAATG | AGTTTAGTAT | TGGAAGTCCA | ATATTCTGAT | TCTCTGTACT | CTTTTCTCT | 2460 |
| ATAAGGCTGA | CAAAAAACAA | ATACAGGAAT | ATGAGCTTCT | ATAGATACAT | CCTCAAAATC | 2520 |
| TTCTCTAGTA | ATCTCTCCAG | ATATCTGAAA | TTCTGGATCT | TGATTTTCCA | ACTCTAAGCC | 2580 |
| TTTTTCTTGC | ATTAATTCCC | AGATTTTTTT | ATTCGTTTCA | GGACTAAATG | TTGCTTGAGT | 2640 |
| TAAGTTCTTA | AAATAAAGTT | CAGGACCACA | CTCGTCAATC | AGCCTCATCT | GCTCTTCCAT | 2700 |

865

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| TTCTGGATAA | GGATTTTCTG | AAAAATCAGC | AAACATGACT | TTTTTAGTTG | TCGGTTCAAT | 2760 |
| TGCTACTAAA | GTCTGACGCT | TAATTGGTTT | CTCGAGTAAT | TTGCAAGCTA | AAATCCACT | 2820 |
| CCAACATATG | GCACAAAGTA | TATATTCAGA | AATTCCTAAT | TCTTCAAGTA | CTTCATAAAC | 2880 |
| CGCATCTGCA | AGATTATCTA | GATTTTTTCC | AGCTTGGTCA | TGAATCGGAC | TCCTACCTGT | 2940 |
| GTTCGGAAAA | TCAATTGTCA | AATAACCAAT | TGTAGGAGGA | GGTTTTTCAA | GTATAAGTGA | 3000 |
| AAAATTTTCA | TAACTTGGTA | GCAACCTGC | TCCGTTTAAA | CAAACTAGCA | CTTCTTTTG | 3060 |
| CTTTTGATAA | GTAACAGAGA | GGCTACCAAT | TTCTGTAGAT | ACTTCAAACC | TCTTCATAAA | 3120 |
| GAAATCCACT | GATTCTATAT | AATGAATTAT | TAAAAATCCT | TATCCTTTAT | TTTATCACGT | 3180 |
| TCCAAGGATT | TTCTCAAGTT | GGAGGAAGGG | GACAATATCT | CTACTTTCCC | TTCAATAATC | 3240 |
| CTTCCAAATT | ATGTTTATGT | TGGTAATTAA | TGGCTGCGGT | TTTGTCTTTC | TCAAAGACAG | 3300 |
| TCTTGGTAA | GTCAATATGA | TTAATAGCTA | CGATTGCGAC | GGTGTAGTAA | ATGATATCAG | 3360 |
| CCAGTTCTCT | GGCAAGTTCC | TCGTCGAAT | CCTATCCCTT | CTTTTCGACC | AGAGCGCCTA | 3420 |
| TTCAAACCT | CGACTACTTC | TCCGACTTCC | TCCACTAACT | TCATAAAGAG | ACCTTCATCA | 3480 |
| GTCCGAGACT | GCTGTTAATG | TTCGATTAA | GATGCTTGA | ATTGCCTAAA | CGTTCAATCT | 3540 |
| TTTATAGTAT | ATTGAACTA | GAATAGTACA | CCTTTACTTC | TAAACATTG | TTAGAAATCG | 3600 |
| ATTTGACTGT | CCTGATCGAT | TTGTCTGTT | CTTGTTCAT | TTTACTATAT | CTTCTATTCC | 3660 |
| ACACAAAAAA | GCGAGACATC | CGTCCCGCCC | TTCTTATTTT | TCGTCAATAA | CGATTCTTAC | 3720 |
| TTTTTTGTAT | TCAGTTGGGA | CAGAGTAGAC | AATCGTTCCT | ATCGCAGAAA | TAGTGCAGAC | 3780 |
| CTTACGACCG | ATTACACGAC | CCACATCGCT | TTGATCAAGA | TTCAAATGAT | ATTCCAAAAA | 3840 |
| TTCTGGTGTA | TCCTCAATCT | TGATAGTTAA | GGCATCTGGT | TGTGAAATTA | AGGGTTTCAC | 3900 |
| AATCGCAATA | ATGAGATTTT | CAATCGTATC | CATCTGTCAA | CCTACTTTAA | ACTTATTTTG | 3960 |
| AAAATTTAGA | ATCGTGGAAT | TTTTTCAATA | CGCCTTCTTT | TGAAAGGATG | TTACGTACTG | 4020 |
| TGTCTGAAGG | TTGAGCTCCA | TTAGCCAACC | ATGCAAGAAC | GCGGTCTTCT | TTCAAAGTTA | 4080 |
| CTTGGTTTTC | AGCAACAAGT | GGGTTGTAAG | TTCCAACCTG | TTCGATGAAA | CGTCCGTCAC | 4140 |
| GTGGTGAACG | TGAATCTGCT | ACGTTGATAC | GGTAGAAAGG | TTTTTTCTTA | GAACCCATAC | 4200 |
| GAGTCAAACG | GATTTTAACT | GCCATTTTAA | AAGTCTCATT | TCTTTAATTT | TTTATTTTCG | 4260 |
| TGAAATAGCT | GAGCTATTTA | GCACATGTTT | TATTATAGCA | GATTTCTGGC | ATGTGTC | 4317 |

(2) INFORMATION FOR SEQ ID NO: 125:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4881 base pairs

866

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 125:

| | |
|---|------|
| AATTTATTTG ACTGGAAATT GTAGAGGGTT CTCGAAATTT CTTGAATGGT TAAAATAAGG | 60 |
| ACAAGAGAAA ACATGGATAT CTATATCCTT GTGCCAAAAA AACCCTGCC CTCCCAGAC | 120 |
| CAACCTGAGG AAAGCAGTGA TTCTTATTTT AGGAGTTAGG AATGAATACA CGAAATCAAT | 180 |
| TTAGCTGATT ATTTTTTGTT TTTCAAGAAT TCATCGTATT GTTTTTGCAT TTCGTTCAAT | 240 |
| ACTTTTTCGT AGGCACCTTC AGATTTCAAT TTTTCCATCA ATTCTGGAAT CGCTTTATCT | 300 |
| GGGTCTACAG TACCACTGTT GATAGCTGTA TCAAATTGTT GCATTGTGTT AGCAATAGCT | 360 |
| GAGATTTTCA GATTTACATT GTCAGTATTG AAGATAAATC CAAGCGCTGG AGATTTCTTA | 420 |
| GCTTCTGCCA ATTTCTTCTT AGAATTTTCG ATTTGTTGGT CTGTAACGTT TTCGTTGATG | 480 |
| TAAAGGATCC AGTTGTTACC AGTGTTCAT CCACCCATGT GAGTGTTCCT TTTGTAGCCA | 540 |
| TCAAGAACGC GAACACGGTT TTCTTTACCT TCAATTTTTT CCCAGTTCTT GCCTTCTGGA | 600 |
| CCGTAAACAA GACCGTTCAA GAGTTCTGGG TTCGTATTCA AGAGGTTCAA GATTTCCATT | 660 |
| GATTTTCTTT TGTCTTAGA GTTGTGTTGAG ATGACAAAGT TAGCAACTTG TGTGTTTGG | 720 |
| TTTTTCTTGA TGAAGTTAGT AATTGGTTTG ATTTGGATAT CTTTGTGGC AACACGTGAA | 780 |
| AGCAAGCTGT TACCGTAGTC AGCTGGTCCT ACTGTTTCTT CACGAACGAA CCAAGTATCT | 840 |
| TGTGTAAGGT CAAAGGAAGT ATCGCTTGTT GCGACGCTTT TTGGAATGTA GCCAGCTTCA | 900 |
| TAGAATTTGT GAAGAGTCTT CAAGTGTCTT TTGAAACGAG GCACTTCGTA ACGGTTTACA | 960 |
| ACTTTAGTAG TATCGCCTTC AAGGTCGATA ACGAATGGAA GACCGTTTGC TACTGGGTAG | 1020 |
| TCAAAATTAT CAGATGGGAT GAAAACCTTA CCAATAGCAA ATGGTACTAC GTCTGGAGCT | 1080 |
| TTTTCTTTGA TTTGTTTCAA GACTGGCTCA AGAGTTTCGT AAGAAGTAAC ACCTGAAATA | 1140 |
| TCGATACCAT ATTTAGCAAG GAGAGTTCCG TTGAAGGCAA AGTTTGTAGA TGATGCAACG | 1200 |
| TTGGCTGCAA CTGGAACAGC GTAAATCTTA CCATTTACAG TATTACCCTT GATGTAAGCT | 1260 |
| GGGTCAAGTG CTTTGTAAGG GTCTTTACCT TCTTTTTTGT ACAATTCTGT CAAGTCAGCG | 1320 |
| TAAGCACCTT TTTGAGCATT TACAATATAG TTATCTGCAA AGGCAATATC ATAGTTTTC | 1380 |
| CCAGATGATG TGATAACTGA CATTTTCTTA CCATAGTCAC CCCAGCCAAG GTATTGGATA | 1440 |
| TCCAATTTGG CACCAACTTT TTCTTCAATG ATTTTGTGTT CATTTGCTAA CAATTCATCC | 1500 |
| AAGTTGTCTG GTTGTCAACC GATTTGGTAC ATTTTGATAA CAGGTTTGTC ACCTGAATCA | 1560 |

867

| | | | | | | |
|------------|------------|------------|-------------|------------|-------------|------|
| GCAGCTTTT | TGCTGTTACC | TGTCAAATTT | CCACAAGCAG | CAAGACCTGC | AGCCAGAGCG | 1620 |
| ACTACACTAG | CAGATGCAAA | AGCATATTTT | TTCCAGTTT | TCATGATAAA | AACTCCTTTT | 1680 |
| TTTATTTT | AACTTATAAA | CAATGTAATG | ATCTTATACT | CAATAAAAAT | CAAAGAGCAA | 1740 |
| ACTAGAAAAC | TAGCCGCAGG | CTGCTCAAAG | CACTGCTTTG | AGGTTGTAGA | TAAGACTGAC | 1800 |
| GAAGTCAGTT | ACATATATCT | ACGGCAAGGC | GACGTTGACG | CGGTTTGAAT | TTGATTTTCG | 1860 |
| AAGAGTATTA | ACTTCACACA | AGGGAAGTTG | GGAAGTGAAG | AATGTTATTT | CTCAATAAGC | 1920 |
| ACTATCTCTT | CACACCACCG | ATAGTCAAAC | CTTTTACAAA | GTAGCGTTGG | AAAAATGGAT | 1980 |
| ACAAAATCGC | GATTGGAAGG | GTTGCAACCA | CAACCATGGC | CATACGACCT | GTTTCTTTTCG | 2040 |
| GTAGAGCAAC | TCCCAGTTGA | CCAATCAAGC | CGACCGCTTT | GGCAATGTAG | TCCATATTTT | 2100 |
| GTTGGATTG | CATGAGCAAA | TATTGCAATG | GATACAAGTT | GTCACCTCTG | ATGTAAAGAA | 2160 |
| GGGCGTTGAA | CCAGTCATTC | CAGAAACCAA | GAGCTGTTAA | GAGCGTGATG | GTTGCGATAC | 2220 |
| CTGGTAGTGA | CAATGGCAAA | CAGATTTGGA | AGAAAATCCG | GGCCTCACTG | GCACCATCGA | 2280 |
| TACGAGCCGA | TTCTAGAATG | GCTTCTGGAA | TGGTCTTCTT | GAAGAAGGAA | CGCATCAAGA | 2340 |
| TGATGTTAAA | TGGTGAGAGA | AGCATTGGA | CAATCAAGGC | CCAAACAGTG | TCACCAAGCT | 2400 |
| GAAGTACACG | GGTCACCATG | ATATAACCTG | GTACCAAAACC | AGCGTTGAAC | AACATACTGA | 2460 |
| GAAGGACGAA | GATGGTAAAG | AATCTGCGAT | ACTTAAAGGT | TGTCCGTGAA | ATAGCGTAGG | 2520 |
| CATAGGTTGT | TGTGATAAAG | ACATTTGTCA | ATGTCCCAAC | TACGGTTACA | AAGACAGAGA | 2580 |
| TGAAGAGGGC | TTGTAGGATT | TTATCCTTAA | ACTGTGCCAA | AAACTCAAAA | CCGTCTAAGC | 2640 |
| CAAATTGGGA | TGGGAAGAAG | CTATAGCCGT | ATTGGAGGAG | GCTTTTCTCG | TCTGTCACTG | 2700 |
| AAATAATGAT | AACGAATACA | AAAGGTAGGA | TACAAGAGAG | GGCAATCAAA | CCCGAAATGA | 2760 |
| TACTGAAGAA | GATATCTGCT | TTCTTACTGA | AGGAGTGAAT | GCCGACATTA | TCAATTTT | 2820 |
| CTTTTTTAAT | TTTCTTTT | GCCATATTCT | CCTCCTTTCT | AGAACAAAGC | TGAGTTTGGA | 2880 |
| TCGACTCGTC | TTGCAAGCAA | GTTTGATAGG | ATAACCAGAA | TCAAACCAAC | AACGGATTGG | 2940 |
| TAAAGACCGG | CTGCTGCAGC | CATACCGATA | TCTGCTGTCT | GAGTCAAACC | ATTAAAGACA | 3000 |
| TATACGTCCA | AAACGTTGGT | TACATTGTAA | AGCTGACCAG | CATTGTGTGG | GATTTGATAG | 3060 |
| AAGAGACCGA | AGTCTGCGCG | GAAGATATTT | CCGACTGCAA | GGATGGTCAA | TACAGTTACA | 3120 |
| AGCGGAGTCA | ACTGAGGAAT | GGTTACGTTG | CGAATACGTT | GCCACTTGCT | AGCTCCGTCC | 3180 |
| ACTGTGCGTG | CTTCGTAGTA | GGTTGGATCA | ATTCCCATGA | TCGTGCGATA | GTACATGACA | 3240 |
| CTGCTATATC | CAAAGCCTTT | CCAAATACCT | AGGAAAAGTA | GGAGATAGGG | CCAGATGCCC | 3300 |

868

| | |
|--|------|
| AGGTCAGCGT AGAAATTGAC TTCTTTGAGA CCAAGACTTT CCAATAGATG ATTGAACACC | 3360 |
| CCTTTATCAA TATTTAGGAA GGCATCTGTA AAGAACTGA TGATAACCCA AGACAAGAAG | 3420 |
| TAAGGGAACA ACATAGAAGT TTGAAAAATC TTCACCATTC TCTTAGAACG GAGCTCGCTG | 3480 |
| AGGATAATGG CAATCCCTAC AGATACAACT AAACCTAGAA AGATAAAGCC AAGATTGTAG | 3540 |
| AGGACAGTAT TTCGTGTGAT AATAAAGGCG TCTCTGAAC TAAATAAGAA TCTAAAATTA | 3600 |
| TCGAGTCCGA CCCATTTACT ATTTATGATA CTATCTATGA AACCATTACT GGTCATGTGG | 3660 |
| TAGTCTTTGA AGGCAACCAC GTTCCCAAAT ACTGGAATGT AAAAGAATAG AATCAACCAG | 3720 |
| AGTGCCCTTG GCAAAACCAT CAAGAGAAAG ATCCAGTTGT CTCTCAATGT TTTTGAAAAC | 3780 |
| TTTTTCATAA TTTCTCCCT TTTTATTTTG ATATCCATCT AAAAATTCCT TTTTAGACTT | 3840 |
| TTGATAACGA TTACATTATT AGTATACTCC TATTTGCAGG TTAGGTTAAA CTCCTAATTA | 3900 |
| TAGAAAAAAC TCCACAAATT ATGTAGCAGA TTTAAACTT TATCACCCT ATCAAACAAA | 3960 |
| TGTCCTAAAT CAATTGTTTA TTTTATCTCT ATTAGCCAG TGATGGCGTC ACTCTGTTAT | 4020 |
| AAGCATCCAA CAACGGGGTA TACTGAAAA TCTCCAGACT AGGGAACCTCA GCGATAGTTC | 4080 |
| CTAATCTGGA GATTTTAAAT ATGTTATTAG GCGTTTGCTT TCAACTTAGC AATAACCTCT | 4140 |
| TTAAGATTAT CAATCAACTC TGCTGCAGTA TGCTCAGAGC CTTTTCATC TGCCAAGAAC | 4200 |
| AAAAGTCTT TTTGAAGTTC TTTTGAGAG TTTTCAAGGA CATCCTTATC TACTGTTTCA | 4260 |
| AGGTTTGAGT CTTAAGAAG TTTACTTAAT TCCTTGCGTA ATTTCTTGAG TTTGATTGTC | 4320 |
| AGACTCATCT TCTCCTGCTG TTTCTTGCC CGCTGTTGT CCTCCATCCT TAGTTGCTGA | 4380 |
| CTGGCTTTCC TTAATGGACT CTAGGGAAGC AATGGCATCT TTGACTGTTT GCAAGATATC | 4440 |
| ACGTAAACCT TGCTCTGTCA AACTATCATC TGCAAAAGCT TTATTAGCCT CTGCCAAAAC | 4500 |
| CAGACGTGCT GAATCTGTGG TAGGATTCGA TACACCTGTC AATGATCTCA AAAGATTTTC | 4560 |
| TAAGGTTTGA GTCTGCTTAC TAATACTAGA CTAAAAATCAA AAAGTATTAT ATAACAGTGA | 4620 |
| TATGAAATCA ACTAAAGAAG AAATCCAAAC CATCAAAACA CTTTAAAAAG ACTCTCGTAC | 4680 |
| AGCTAAATAT CATAAACGCC TTCAAATCGT TCTATTTTGT CTGATGGGCA AATCTTATAA | 4740 |
| AGAGATTATA GAACTTTTAT AGTAGTTTGA AATAAGATGT GAACATCTCT ATCAGGAAAG | 4800 |
| TCAAATTAAT TTATAGAAAT ATTTTAGCAG CCAAGGTGTA CTGTTATAGA TTCAATACAC | 4860 |
| TATACTTGGT GGTTTAGCTC G | 4881 |

(2) INFORMATION FOR SEQ ID NO: 126:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 13121 base pairs
 - (B) TYPE: nucleic acid

869

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 126:

| | |
|---|------|
| AGGATCCCCG GAAAAGGAGA CTAAAAATGA AGAAAAAATT TCTAGCATTT TTGCTAATTT | 60 |
| TATTCCCAAT TTTCTCATTA GGTATTGCCA AAGCAGAAAC GATTAAGATT GTTCTGATA | 120 |
| CCGCCTATGC ACCTTTTGAG TTTAAAGATT CAGATCAAAC TTATAAAGGA ATTGATGTTG | 180 |
| ACATTATTAA CAAAGTCGCT GAGATTAAAG GCTGGAACAT TCAGATGTCC TATCCTGGAT | 240 |
| TTGACGCAGC AGTCAATGCG GTTCAAGCTG GGCAAGCCGA CGCTATCATG GCAGGGATGA | 300 |
| CAAAGACTAA AGAACGTGAA AAAGTCTTCA CCATGTCTGA TACTTACTAT GATACAAAAG | 360 |
| TTGTCAATTG TACTACAAAG TCACACAAAA TTAGCAAGTA CGACCAATTA ACTGGCAAAA | 420 |
| CCGTGGGTGT TAAAAACGGA ACTGCCGCTC AACGTTTCCT TGAAACAATC AAAGATAAAT | 480 |
| ACGGCTTTAC TATTAAAACA TTTGACACTG GTGATTTAAT GAACAACAGC TTGAGTGCTG | 540 |
| GTGCCATCGA TGCCATGATG GATGACAAAC CTGTTATCGA ATATGCCATT AACCAAGGTC | 600 |
| AAGACCTCCA TATTGAAATG GATGGTGAAG CTGTAGGAAG TTTTGCTTTC GGTGTGAAAA | 660 |
| AAGGAAGTAA ATACGAGCAC CTGGTACTG AATTTAACCA AGCCTTGCTCT GAAATGAAAA | 720 |
| AAGATGGTAG TCTTGATAAA ATTATCAAGA AATGGACTGC TTCATCATCT TCAGCAGTGC | 780 |
| CAACTACAAC TACTCTCGCA GGATTAAAAG CTATTCTGT TAAGGCTAAA TATATCATTG | 840 |
| CCAGCGATTG TTTCTTGCC CCTTTGTTT TCCAAAATC AAGCAACCAA TACACTGGTA | 900 |
| TTGATATGGA ATTGATTAAG GCAATCGCTA AAGACCAAGG TTTTGAAATT GAAATCACCA | 960 |
| ACCCTGGTTT TGATGCTGCT ATCAGTGCTG TCCAAGCTGG TCAAGCCGAT GGTATCATCG | 1020 |
| CTGGTATGTC TGTACAGAT GCTCGTAAGG CAACTTTTGA CTTCTCAGAA TCATACTACA | 1080 |
| CTGCTAATAC CATTCCTGGT GTCAAAGAAT CAAGCAATAT TGCTTCTTAT GAAGATCTAA | 1140 |
| AAGGAAAGAC AGTCGGTGTT AAAAACGGAA CTGCTTCTCA AACCTTCCTA ACAGAAAATC | 1200 |
| AAAGCAAATA CGGCTACAAA ATCAAAACCT TTGCTGATGG TTCCTCAATG TATGACAGTT | 1260 |
| TAAACACTGG TGCCATTGAT GCCGTTATGG ATGATGAACC TGTCTCAAA TATTCTATCA | 1320 |
| GCCAAGGTCA AAAATTGAAA ACTCCAATCT CTGGAACCTC AATCGGTGAA ACAGCCTTTG | 1380 |
| CCGTTAAAAA AGGAGCAAAT CCAGAACTGA TTGAAATGTT CAACAACGGA CTTGCAAACC | 1440 |
| TTAAAGCAAA CGGTGAATTC CAAAAGATTC TTGACAAATA CCTAGCTAGC GAATCTTCAA | 1500 |
| CTGCTTCAAC AAGTACTGTT GACGAAACAA CGCTCTGGGG CTTGCTTCAA AACAACTACA | 1560 |

870

| | |
|--|------|
| AACAACTCCT TAGCGGTCTT GGTATCACTC TTGCTCTAGC TCTTATCTCA TTTGCTATTG | 1620 |
| CCATTGTCAT CGGAATTATC TTCGGTATGT TTAGCGTTAG CCCATACAAA TCTCTTCGCG | 1680 |
| TCATCTCTGA GATTTTCGTT GACGTTATTC GTGGTATTCC ATTGATGATT CTTGCAGCCT | 1740 |
| TCATCTTCTG GGAATTCCTA AACTTCATCG AGTCTATCAC AGGCCAACAA AGCCCAATTA | 1800 |
| ACGACTTTGT AGCTGGAACC ATTGCCCTCT CACTCAATGC GGCTGCTTAT ATCGCTGAAA | 1860 |
| TCGTTCTGGG TGGTATTCTG GCCGTTCCAG TTGGCCAAAT GGAAGCCAGC CGAAGCTTGG | 1920 |
| GTATCTCTTA TGGAAAAACC ATGCGTAAGA TTATCTTGCC ACAAGCAACT AAATTGATGT | 1980 |
| TGCCAAACTT TGTCAACCAA TTCGTTATCG CTCTTAAAGA TACAACTATC GTATCTGCTA | 2040 |
| TCGGTTTGGT TGAACCTCTC CAACTGGTA AGATTATCAT TGCTCGTAAC TACCAAAGTT | 2100 |
| TCAAGATGTA TGCAATCCTT GCTATCTTCT ATCTTGTAAT TATCACACTT TTGACTAGAC | 2160 |
| TAGCGAAACG CTTAGAAAAG AGGATTTCGT AATGGCAAAA TTAATAATTG ATGTAAATGA | 2220 |
| TTTACACAAG CACTATGGAA AAAATGAAGT CCTAAAAGGA ATTACGACTA AGTTCATGA | 2280 |
| AGGAGATGTT GTTTGTATCA TCGGTCCTTC AGGTTCTGGT AAGTCAACTT TCCTCCGTAG | 2340 |
| CCTCAATCTT TTAGAAGAAG TCACTAGCGG TCACATCACT GTGAACGGCT ATGATTTAAC | 2400 |
| TGAAAAAACA ACCAATGTTG ACCACGTCCG TGAAAATATC GGCATGGTAT TCCAACACTT | 2460 |
| CAACCTCTTC CCTCATATGT CTGTATTGGA CAACATCACC TTTGCTCCTA TTGAGCACAA | 2520 |
| GTGATGACT AAGGAAGAAG CTGAGGAATT GGAATGGAG TTGCTTGAAA AGGTGGACT | 2580 |
| AGCAGATAAA GCTAATGCCA ATCCAGATAG CCTATCAGGT GGTCAAAAAC AACGTGTGGC | 2640 |
| CATCGCTCGT GGCCTAGCAA TGAATCCAGA CATCATGCTC TTCGATGAAC CAACTTCTGC | 2700 |
| CCTTGACCCCT GAGATGGTTG GAGACGTACT TAACGTTATG AAGGAATTGG CTGAGCAAGG | 2760 |
| CATGACCATG ATTATCGTAA CCCATGAGAT GGGATTTGCT CGTCAGGTTG CCAACCGCGT | 2820 |
| TATCTTTACT GCAGATGGCG AGTTCCTTGA AGACGGAACA CCTGACCAA TCTTTGATAA | 2880 |
| CCCACAACAC CCTCGTCTGA AAGAGTTCTT AGATAAGGTC TTAAACGTCT AAACCAAAC | 2940 |
| TGTAAGGATT TCCTTGCACT TTTTCTACCT CGTATTGGAA TTTTGTATT TTCGGAAAAT | 3000 |
| TATGTTAGAA TTAAGTTTAT GAAATGAGGT TTCCTCATAC CTAGCAAGAC TAGGAATAAA | 3060 |
| AATAGAAATT AGGTAGCTAG ATGTCATCTA AGGTTATTGT TACAATTTTC GGTGCGAGTG | 3120 |
| GAGACCTGGC TAAACGCAAG CTCTACCCCT CCCTTTTCTAG ACTATATCAA TCCGGCAATC | 3180 |
| TTTCCAAGCA CTTTGCCGTT ATTGGAACG CCCGTAGACC TTGGAGTAAG GAATATTTTG | 3240 |
| AATCTGTAGT TGTCGAGTCC ATCCTTGATT TGGCAGATAG TACCGAGCAA GCCCAAGAAT | 3300 |
| TTGCTAGCCA CTTCTACTAT CAAAGCCATG ATGTCAATGA TTCGGAACAT TATATTGCTT | 3360 |

871

| | |
|--|------|
| TGCGTCAATT ACAAGCTGAG CTTAATGAAA AATACCAAGC TGAACACAAT AAGCTCTTCT | 3420 |
| TCTTGTCTAT GGCACCTCAG TTCTTTGGAA CCATTGCCAA ACACCTCAAA TCTGAAAACA | 3480 |
| TTGTCGATGG CAAAGGTTTT GAGCGCTTGA TCGTTGAAAA ACCATTTGGT ACAGATTACG | 3540 |
| CAACTGCAAG CAAGTTGAAT GACGAACTCC TAGCAACATT TGACGAAGAA CAAATTTTCC | 3600 |
| GTATCGACCA TTATCTTGGT AAGGAAATGA TCCAAAGCAT CTTTGCAGTT CGCTTTGCAA | 3660 |
| ACTTGATTTT TGAAAACGTT TGGAAACAAGG ATTTTATCGA CAATGTTCAA ATTACCTTTG | 3720 |
| CGGAGCGCTT GGGTGTAGAA GAACGTGGTG GCTACTATGA CCAATCCGGT GCCCTCCGTG | 3780 |
| ACATGGTCCA AAACCACACT CTACAACCTC TTTCGCTCCT CGCCATGGAC AAACCAGCAA | 3840 |
| GCTTCACAAA AGACGAGATT CGTGCTGAAA AGATTAAGGT CTTTAAAAAC CTCTATCATC | 3900 |
| CAACTGATGA AGAACTCAAA GAACACTTTA TCCGTGGGCA ATACCGCTCT GGTAAGATTG | 3960 |
| ATGGCATGAA ATACATCTCT TATCGTAGCG AGCCAAATGT GAATCCAGAA TCAACAACTG | 4020 |
| AAACCTTTAC ATCTGGTGCC TTCTTTGTAG ACAGCGATCG ATTCCGTGGT GTTCCTTTCT | 4080 |
| TTTTCCTGAC AGGTAAACGA CTGACTGAAA AAGGAACTCA TGTCAACATC GTCTTTAAAC | 4140 |
| AAATGGATTC TATCTTTGGA GAACCACTTG CTCCAAATAT TTTGACCATC TATATTCAAC | 4200 |
| CAACAGAAGG CTTCTCTCTT AGCCTAAATG GGAAGCAAGT AGGAGAAGAA TTAACTTTGG | 4260 |
| CTCCTAACTC ACTTGATTAC CGTACAGATG CGACTGCAAC TGGTGCTTCT CCAGAACCAT | 4320 |
| ACGAAAAATT GATTTATGAT GTCCTAAATA ACAACTCAAC TAACTTTAGC CACTGGGATG | 4380 |
| AAGTTTGTGC GTCATGGAAG TTGATTGACC GTATTGAAAA GCTCTGGGCT GAAAAAGGTG | 4440 |
| CCCCACTTCA TGA CTATAAA GCTGGAAGCA TGGGACCTCA AGCCAGCTTT GACCTACTTG | 4500 |
| AAAAATTCGG TGCCAAATGG ACTTGGCAAC CAGATATCAC CTATCGTCAA GATGGTCGCT | 4560 |
| TAGAATAAAA AAATTTCTTG CAAGTTTATG CcTTGCAGGA TTTTGTCTTC TGATTAGATT | 4620 |
| AAACCTTCCA AGAGACCTTT CATAAAGTTT TCTGAGTTAA ACTCTCCAAT ATCATCGATT | 4680 |
| TTTTCACCAA AACCAATCAA TTTTACAGGA ATATTGAGTT CTTACGAAT GGCTAGAACC | 4740 |
| ACACCTCCTC GAGCAGTTCC ATCAATCTTA GTCAAAACAA TTCCCGTTAA AGGTGTGATT | 4800 |
| TTCGAAAAAT CTTTGGCCTG TACTAGGGCA TTTTGACCTG TTGATGCATC AAGTGCCAAG | 4860 |
| AAGGTTTCAT GTGGTGCTTC TGGCACAACA CGTTTGATAA TACGACCAAT CTTTCCAAC | 4920 |
| TCAGCCATAA GGTATCTCTT ATTTTGAGCA CGACCAGCAG TATCAATCAT GAGAATATCG | 4980 |
| ATACCTTCAG TCACGGCAGG TTCCATACCA TCAAAGACCA CGCTGGCTGG ATCAGCTTTT | 5040 |
| TCAGGTCCAG TTA CTACTGG AACATCTACT CGTCGGCCCC ATTCACTAG CTGAGCTACT | 5100 |

872

| | |
|--|------|
| GCACCCGCAC GGAAGGTATC TGCTGCAACC AGCATGACCT TCTTACCAGC TTGTTTGTAG | 5160 |
| CGGTGGGCTA GTTTTCCGAT AGAAGTTGTT TTCCCAACAC CATTACACACC AACAAAGAGC | 5220 |
| ATAACTGTCA AGTTATCTTG GAAGTGGATG CTTTCATCGT AGCTACCATC CTTTTCATAA | 5280 |
| AGCTCAACCA ATTTCTCAAT GATGACACGA CGAAGTACAT CAGGTTTCTT GGCATTTTCA | 5340 |
| AGCTTGGCTT CGTAACGTAG TTCTCCGTT AAGTTAGAAG CGACTTGGAC ACCAACATCA | 5400 |
| CTCATAATCA GCAGTTCCTC CAGTTCCTCG AAAAAATTCTT CGTCAACAGA GCGGAAGTTA | 5460 |
| GCAAAGAAGG CATTCAAGCG GGCACCGAAA CCTGTGCGAG TTTTCTTAAG ACTGCGGTCA | 5520 |
| TATTTTCTCT GAACAGTTTC TTCTGTTTGA GGAGCTTCTG GTTCAAGCAC TTCAGAATTA | 5580 |
| TTTTCTTCTA CAGTTCCTTC GTGCTCAAGC TTCTCTTCCT CTGGTAATTC TTCTGAGTTT | 5640 |
| GGTAATTCCT CTATTTCTTC TTGAGAAACC CCTACAGCTG GCTCTGAATC CTGACTTTCT | 5700 |
| TCAACTGTGT CTTGGATTTC CTCTTCTTGG AACACAGCTT GTTCAACAAT TTCAACCTCT | 5760 |
| GCTTCTTCCT GAGAACTTC TCAACTTCT GTGAAGGTAG GATCAACATC TTCAGACAAA | 5820 |
| TCAAGATTTT CCAGAGCTTC TTTTACAAC TCTTCGATT TAGGTTCTTC TTTTTTCCG | 5880 |
| AATAGACGGT CAAACAATCC CATATCTTAG TTCTCCTTTA GCACATATTC TTCGATAGCC | 5940 |
| CAGGCGACAG CTTCTCATC GTTGGTCATC GCGGCTACTA CATTTGCGGC TGCTTTTACT | 6000 |
| TCAGGAACAG CGTTTTGCAT AGCAACACCA AGACCTGCCC ATTCAATCAT AGAGAGGTCA | 6060 |
| TTGGCTCGT CACCACAAGC CATCACTTGA CTTTGGTCGA TTCCAAGATG GCTGATTAGT | 6120 |
| TTTGCCAAAC CTGTTGCTTT ATGAACATTC TTTGGTGACC ATTCTAGCAA CATTTACGCT | 6180 |
| GATTTAAAGA TTTCATATTG GTCAAACAAT TCTGGAGAAA TCTTCTGAAT GGCTGCATCC | 6240 |
| AAGGGTTCTT GAGCAAAGGC AGTCACGCAT TTGTTGTAGG TCATTTGACT AGATAAGTCT | 6300 |
| TCAAAGTCCA CTGGAACAAA GGTCAAAGCT GGATTGAATT TGGCATAAAG ACTTTCTTGG | 6360 |
| TCCGATTGGA TTGATAAAC GTTTCCTTCT GAGATGGCAT CAAGAGGCAG TGATAATTC | 6420 |
| TCTGTTTCTT CATACAAACG TGCCACATCA TCATATGAAA AGACTGTTTT ATCAAGGATT | 6480 |
| TCTCCTGTAT TTTTCTGAAC TAATCCACCA TTAAGGTAA TGGTATACTC ATCTTCCTGA | 6540 |
| CCGTCAGTCC CTAACATATG GAGAAAGAAA TCCATGGCTT TTAAGGGACG ACCAGTTGTC | 6600 |
| AATACGACCT TGATACCAGC ATCAGCGCA gCTTGCAAGG TTTCTTGGT ACATCCGTC | 6660 |
| AGCCTTTTAT CAGTAGTCAG CAAGGTCCCG TCCAAGTCCA ATGCAATCAA TTTTATATCT | 6720 |
| GCCATTATAA GCCCTCCATA TAAGCTATAA CCGACCGTTC CTTATGGTGA CCAATCACAG | 6780 |
| TCTTTGCTAA TTCTAAATTT TCAGGTCGTG CATTTTCAGG AGCTACAGGA TGTCCCACAA | 6840 |
| CCTGCATCAT ATGTAAGTCA TTAAGATTGT CTCCAAAAGC CATGACCTGA TCCATTGTGA | 6900 |

873

| | |
|--|------|
| TACCAAGTTT TTAACTAAT TCAACAATGG CCACTCCCTT ATCGACATAG TCCAGAACAA | 6960 |
| TATCAATGGA TTCAAAGCCA GTTGTCATGG CCTTAACACC AGGAACGTTT TCGTTTACCC | 7020 |
| AAGCCTCCCC ATCTTCCAGC GTTTCTTCTG TGAAGTTGGT TGTAAATTG AAAATGTCAT | 7080 |
| CTGTGATATC TTCCAAACTC GCTACTTTT TTGGATATTTT ATTATAGTGC TGA CTCACTT | 7140 |
| TCAAATAGGT CTCATCAACC GTATCTAGAA CATATGAACC CTTCTTACCC GTCAAGAGCA | 7200 |
| GTTTATTGAT ATCTACATAA GGTGAAGTTT TCAGCTTTT AAAAGTTGCC AGATAAAAGT | 7260 |
| CACGAGACAT AGTCGCTTCA TACAAGTCCT GACCTTGATA CTCTACCAA CTGCCATTTT | 7320 |
| CCGCGATGAA AATAATGTCA TCACGAACAC CAGCAAATAA TTTTCTAGA GACAGAAATC | 7380 |
| CCCGACCCGA AGCTACCGCA AAGTAAATCC CTTTTCCTT GTAGGAAACC AAGAGAGACT | 7440 |
| TGAGACGATC CATATCAAAG CGTCCATTCC CATCTAGGAA GGTTCCTGCC ATATCCGTTG | 7500 |
| CTACTAGTTT AATTGTCATC CTTCAACTT TCTAAATCT TTTAACTTAA CTGAAACAAT | 7560 |
| CTTTGAAACA CCGATTCTT GCATGGTCAC TCCATAGATG GAATCAGCCG CTGCCATGGT | 7620 |
| TCCCTTACGG TGGGTACGA CGATGAACTG GCTGTCTTG TCAAAGCGGT TGAGGTAATC | 7680 |
| CCCAAACGT TTAACATTGG CTTTATCCAG CGCAGCTTCC ACCTCATCCA AGATAACAAA | 7740 |
| TGGAATAGTC TTGACACGAA TAATGGAGAA GAGCAAGGCA AGAGCCGATA GGGCTTTTTC | 7800 |
| ACCACCACTC ATGAGATTAA GAGACTGGAT TTTCTTGCCT GGTGGTTGGA CAGAAATTC | 7860 |
| AACCCAGCT GTCAGCAAGT CTCCTTCAGT CAAAATGAGG TCAGCCTGAC CTCCACCAAA | 7920 |
| CATCTGCTTG AAGGTCACTT TAAAGGACTC ACGAATGACC TCAAAGGTTG ATTTAAAGCG | 7980 |
| TTCTTGAACC TCATCATTCA TCTCTGTAAT GGTCTCAAGG AGCAGGTTT TCGCAGACAA | 8040 |
| AATATCATCA CGTTGGCTAT TTAGGAAATC CAGACGGTTG TGAACCTCTT CGTACTGTTT | 8100 |
| AATAGCGTCT AAATTGACAG GACCCAGTGA GCGTATAGCC TTCTCTAAAT CCTTAACTTC | 8160 |
| TTGCTCTGCC AGATTGAGAT TTTCCAATC ATGCGCCTT TCTAAAGCTT CTGTGTAGCT | 8220 |
| GATCTGGTAC TGGTCTGTTA ATTGACTTTG TAGATGGCGC AAGCGCTCGC TAACCTTTTC | 8280 |
| TTTCTTGGCT TCAGCACGAG TTTGCTTGCG AATCCACTCT TCATTCTGCT GCGGAGCCTG | 8340 |
| ATCCAAATGA CTAGCAATAT CATCCAGTTG ACCCTCAATA TCATCCAACT CAACTGCTT | 8400 |
| GCGAATCAAA CCTTGTGGA GATTGTTTT TTGAGTTTTG GATTCTTCG CCTGTTGACT | 8460 |
| GAGCAATTCT GTATCAACCT TCTCAAGATT ATCAATCTTT TCTTGAAGAA GCGCTGGAT | 8520 |
| TTCTCTTGT TCAAATCAA GATTGTCCAA TTCTTGCCT AAGCGTTCAA TATCAGCAAC | 8580 |
| TTCATAACGT TTTTGCCCTT GCAGTTCTGT CTTAAGCAAA CGAGCTTGCG CTAGCTCTTC | 8640 |

874

| | | | | | | |
|-------------|------------|------------|-------------|------------|------------|-------|
| CTGCAAGTTT | TGATAGCGTT | CTTGGATGGC | ATTTTGTGTTA | GACTTAATCT | CTTCAATCTC | 8700 |
| AGCTTCCAGA | TTTTGCTTGT | CACTGGAGAT | TGCAGCAAGA | CGCTCTTGGC | AGTTTTCCTT | 8760 |
| ATCCGCTTGC | CAATCTCCCT | CGGAAAGACG | ATCTATTTCC | TCTTCTTGGA | GTTTCCAAAG | 8820 |
| AGTTTCCAGT | TCTTCAACTT | GCTGACTAGT | TTGCTGATAA | GCGAGGAACA | AGCCTTGCTC | 8880 |
| CTGAATACGT | GCCTGCTCTC | CTTGAGATTT | AATAGCTTCT | AATGACTCGG | TCAATCTGGC | 8940 |
| CATCTCATCT | TGCAAGGTCT | TCAAAGTCGC | CTCTTCTGAA | CCCAAGCTTG | CTTCTTCTTC | 9000 |
| AGCAATTTCT | TTTTGTAATT | GCTCCAGTTC | TGGCTTGATA | AAAATGCTGT | TATCTTGGCG | 9060 |
| ATTGGCACCA | CCTGCATAAG | AACCACCTGT | GCGCAACTCT | GTCCCATCCA | ATGTCACCAT | 9120 |
| ACGAACCTGA | TAACGAACCT | GGCGAGCTGC | TGCACGCGCA | TGTTCTACGG | TATCAAAGAT | 9180 |
| AGCCGTCGTA | GCTAGCAAGT | TCTTGAAAAT | GGCTTCCAGT | CTAGTATCAA | AAGTCACCAA | 9240 |
| CTCATCTGCC | ATCCCAAGGA | AACCTGGGCT | TACAGCGATA | GCATCTTGGT | TCTGACTAGA | 9300 |
| AATCGTACGC | GCCTTGATAG | TGGTCAAAGG | AAGAAAGGTT | GCACGACCGG | CTCTGTTCGG | 9360 |
| TTTAAGGAAG | TCAATAGCCT | TGGTTGCCGA | CTCTTCATCT | TCTACGATGA | TATGCTGGCT | 9420 |
| ACTTGCCCCCT | AAGGCAATCT | CTAGGGCAGT | TTGATAATAA | ACATCAAAGG | TCAGATGCTC | 9480 |
| ACTGACTGCA | CCAATAATCC | CACCTAGGCG | ATCTTTTCT | TGGAGAACAC | TCTTAACACC | 9540 |
| TGCATAAAAG | TTACTATGAT | TTCTCAGGAT | ATTTTCCAAA | CTTTGAGCTC | TGGCCTGCCT | 9600 |
| GTTTTTGAGA | TTATCCAGAC | GGTCAAAGAG | TTGGCTTTGT | TGAGCTTGAT | AGGAAGTTTT | 9660 |
| CTGCTCCTCT | TGCTCCTTGG | CAATAGCTTG | GATGTCAGCC | AATAATTTCT | GAACCTGCTC | 9720 |
| CTTGGCAGTT | TCAAGCTCTT | CCTTTTGCTG | ACTAGCCTTC | TCTTTAGCTA | TAGCTAATTG | 9780 |
| CTCTTTCAGC | TTTTCTAGTT | GATCTGCTTG | TTTTTGAGAA | AGCTGACGAC | TATTTTCCAA | 9840 |
| CTCATTTCTCA | ATACGGGTCA | ACTGGTTTGA | GACATCCGCT | TCTTCTTGTA | AAAGAGCTAC | 9900 |
| AAAGCGTTCA | CGTAAGAGCT | CAATCATCTG | ATCAGGATCG | TCTGAGAAAG | CCAGCAATTC | 9960 |
| AGCTTCTAAA | CGATTGAGTT | TTTGATTATT | TTGGACTAGA | TTTCCCTCTA | ACAGAGCTAA | 10020 |
| AGAGCTTTCT | TTATCAGACT | TTTCTTTGCT | GAGTGAATTT | CTCTTATCCT | CCAAAGCAGC | 10080 |
| CAAACGGGCT | TGTGCCTCCT | GTTGATTCAA | GGCCACTTGC | TCGGACTCCA | GTTTCGATAG | 10140 |
| GGCTAATTTT | CTTTCTAAAT | CACTAATCAG | ACTAGTCAAG | TCCATCAAAC | TGCCTTGGTC | 10200 |
| TTTGGCCATT | TCAGCCTGTA | AATCTTGGCG | TTGCTTTTGA | AGAGTTTGAT | TTTCTTCTTC | 10260 |
| TAATTTTTC | CGCTTTTGGT | AATAACTCAT | CAAGAGTTCT | TGAACCTGAG | TCAACTCTTC | 10320 |
| TTCTGTCGAC | TCTAGTTCAG | CCTTATTTTC | CTTGATTGTA | GCAACCAGAA | CATCTAAATA | 10380 |
| AATAGCCTTA | CGTTGTCCTT | CCAAGTCTAA | AAACTTACGG | GCATTCTCAG | CTTGCTTCTC | 10440 |

875

| | | | | | | |
|------------|------------|------------|-------------|------------|-------------|-------|
| AAGAGGCTTG | ATTTGATTAT | CCAACTCGTA | GATAATGTCC | TCTAAGCGGT | CCAGATTATC | 10500 |
| CTGAGTTTGC | TGCAGTTTAC | TCTCGGTTTC | TTTTCTGCGA | GTCTTGATTT | TTAAAACTCC | 10560 |
| AGCAGCTTCT | TCAAAAATAG | CTCGTCGTTT | CTCAGGCTTG | GAATTAAAAA | TCTCCTCAAC | 10620 |
| CTTCCCTTGG | GAAATAATAG | AGAAGGAATC | TCGTCCCAAT | CCAGTATCCA | AGAAGAGGTC | 10680 |
| ATGAATATCA | CGCAGACGGA | CTTCTTGCC | GTCAATCTTG | TATTCGCTAT | CTCCACTACG | 10740 |
| ATAGACATGG | CGTCCACCC | TGATTTCTTG | ACCTGCATCC | TTGATAAATC | CGTCATGATT | 10800 |
| ATCCAGAGTC | ACAACTACAG | AAGCATAATT | GAGCGGTTTG | CGACTTTCGG | TTCCAGCAAA | 10860 |
| GATGATATCC | GGCATCTTGC | CCCCACGGAG | ACTCTTGACA | CTAGACTCCC | CCAAAGCCCCA | 10920 |
| ACGCAGACTT | TCTGTAATAT | TGGACTTTCC | AGATCCATTG | GGTCCAACAA | CTGCCGTCAC | 10980 |
| ACCTTGGTCA | AAAACGACCT | TGGTCTTATC | AGCAAAAGAC | TTGAACCCCT | GAATTTTCGAT | 11040 |
| TTCTTTTAAA | TACATGAATC | CAGCCCCTTC | TCAACGGCAT | TTTTGGCAGC | TTCTGCTCT | 11100 |
| GCTAATTCT | TAGAACGACC | TTGGCCTTGA | CCGATGCTCT | TACCTTCAAC | AAGAACTTCT | 11160 |
| ACATCAAAAA | CCTTATCGTG | AGCAGGCCCT | GTTCAGAAA | TCACCTGATA | ACGAATAGCC | 11220 |
| ACATCACCAT | TGACCTGAAG | CAACTCTTGG | AGATGGGTTT | TATAGTCTGT | AATCATCTCA | 11280 |
| AACTCGCCTG | CTTCAACCTT | AGGAATCATG | ACTTGATAGA | TAAATTCCTT | GACCTTGGCC | 11340 |
| ACATCCTTAT | CCAAAAGAAG | GGCACCAAGA | AAGGCTTCAA | AGGCATCACC | AAGAATGGTG | 11400 |
| TCACGATTGC | GACCACCTGA | TTTTTCTTCC | CCTTTACCCA | ACTTGATAAA | CTGGTCAAAC | 11460 |
| TGGCAATCAC | GCGCAAAACC | AGCTAAACTC | TCCTCACGGA | CAATCATAGC | ACGGAGTTTT | 11520 |
| GATAGGTCAC | CTTCAGGCTT | TTTAGGATAT | TTTTTATATA | GATATTCTGA | AATCAATAAC | 11580 |
| TGTAGAACAG | CGTCTCCTAA | AAATTCCAAG | CGTTCATTGT | GTGAAATTTT | TAAGAGGCGG | 11640 |
| TGCTCATTTG | CATAACTCGT | ATGAGTAAAG | GCAGTTTCCA | GTAACCTTTT | GTCTGCAAAT | 11700 |
| TCGATTGCAA | AATGATTCTT | TAGTACAGTT | TGTAATTCTT | TCATACCAAC | CTCTTTCTAA | 11760 |
| CTGATAATAG | TCCTTTTTAT | TATATCAAAA | AAAGCCCCCT | GAGTCACTCT | AAAACGGGAC | 11820 |
| TGGAAAGCAT | TTGGGAATTC | TTTAGACAGA | GATTCTCAGT | TTTAGCGGCA | AATTTGGGTC | 11880 |
| AGGATAAAGA | AAAAAGCCCT | ATTAAAGGCT | TTTTAGGATG | TTTACATCCA | CCCTGAGGGA | 11940 |
| ATCGAACCCC | CATCTCAAGA | ACCGGAATCT | TACGTGATAT | CCATTACACT | AAGGGTGGAA | 12000 |
| ACTTGTTTTA | TTATAACAGA | AATTTGCTCT | AATAACAAAGT | TTTTTGGTCA | AAGACCCCGT | 12060 |
| CTTAGTGGGA | AGCATCCCCA | TTCCAGATGG | AGTTTTTCAC | GATCACATAA | TCAACGTGTT | 12120 |
| TAAGGTCAGC | AACCTGACGT | CCACCTGCAT | AAGAAATAGC | ACTTTGAAGG | TCTTGTTCCT | 12180 |

876

| | |
|---|-------|
| TCTCAGTTAA AGTGTCTTGC AGATGACCTT TAGCAGGAAG CAAGATACGT TTGCCTTCCA | 12240 |
| CATTTTTGTGTA AGCACCTTTT TGATATTGTG AGGCTGAACC ATAATATTCT TTGAACTGTT | 12300 |
| CACCATCGAC TTCAATCGTT TTCCCTGGAC TTTCAATGTG TCCTGCAAAG AGGGAACCAA | 12360 |
| TCATGATCAT GCTAGCACCG AAGCGGATAG ACTTAGCAAT ATCACCGTGA GTACGAATTC | 12420 |
| CTCCATCAGC GATAATCGGT TTACGCGCAG CCTTGGCACA CCAGCGTAGA GCAGCCAACT | 12480 |
| GCCAACCACC TGTACCAAAA CCAGTCTTAA CCTTGGTGAT ACAAACCTTA CCAGGACCGA | 12540 |
| TTCCGACCTT AGTAGCATCC GCACCAGCAT TTTCCAATTC ACGCACAGCT TCTGGTGTTT | 12600 |
| CCACATTTCC AGCAATGACA AAGGTATCTG GCAATTCCTT CTTGATGTGT TGAATCATAG | 12660 |
| AAATCACGCT ATCCGCATGA CCATGAGCAA TATCAATAGT GATATACTCA GGAGTATCAG | 12720 |
| CCTTGAGCTG GCTAACAAAA TCATACTCAT AATCCTTAAC ACCGACAGAG ATAGAAGCAA | 12780 |
| TGAGCCCTTG ATTGTGCATT CGTTTAATAA AAGGAATGCG TCCTGCCTCA TCAAAACGGT | 12840 |
| GCATAATGTA GAAGTAACCA CCTTTAGCCA GTTGCTCTGC TACATTTTCA TCCAAAATCG | 12900 |
| TCTGCATATT CGCTGGCACA ACAGGTAGTT TAAAGGTGTG ATTCCTAAA GTGACACTTG | 12960 |
| TATCCGCTTC TGCACGGCTT TTAATGACAC ATTTATTTGG AATCAATTGA ATATCTTCGT | 13020 |
| AATCAAAAAT TGGAAATTC TTTAACATAT CGATGTCTCG TTTCTTTTGT AATGACCTAC | 13080 |
| CTATGCTCTT GCATCACTAC GCCTTTTCCG ACGTTTCCTG G | 13121 |

(2) INFORMATION FOR SEQ ID NO: 127:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 9578 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 127:

| | |
|---|-----|
| CCGAATGCAA TGTTTACGGT TGAACCTGAA AATGGACATC AGATTTTAGC AACAGTTTCT | 60 |
| GGTAAATTC GTAAAACTA TATTCGTATT TTAGCGGGAG ATCGTGTAC TGTCGAAATG | 120 |
| AGTCCATATG ACTTGACACG TGGACGTATC ACTTACCGCT TTAAATAATC GAAAACTTG | 180 |
| GAGGGATAAG AAATGAAAGT AAGACCATCG GTCAAACCAA TTTGCGAATA CTGTAAAGTT | 240 |
| ATTCGTCGTA ATGGTCGTGT TATGGTAATT TGCCCAGCAA ATCCAAAACA CAAACAACGT | 300 |
| CAAGGATAAG ATAGAAAGGA GAAAACATGG CTCGTATTGC TGGAGTTGAT ATTCCAAATG | 360 |
| ACAAACGCGT AGTAATCTCA TTGACTTATG TTTATGGTAT CGGACTTGCA ACATCTAAGA | 420 |
| AAATTTTGGC TGCTGCTGGA ATCTCAGAAG ATGTTCTGTG ACGTGATCTT ACATCAGATC | 480 |

877

| | |
|--|------|
| AAGAAGATGC TATCCGTCGT GAAGTGGATG CAATCAAAGT TGAAGGTGAC CTTCGTCGTG | 540 |
| AAGTAAACTT GAACATCAAA CGTTTGATGG AAATCGGTTC ATACCGTGGT ATCCGTCACC | 600 |
| GTCGTGGACT TCCTGTCCGT GGACAAAACA CTA AAAACAA CGCCCGCACT CGTAAAGGTA | 660 |
| AAGCTGTGTC GATTGCTGGT AAGAAAAAT AATATAGGAG GTAAAAGTCT TGGCTAAACC | 720 |
| AACACGTAAA CGTCGTGTGA AAAAGAATAT CGAATCTGGT ATTGCTCATA TTCACGCTAC | 780 |
| ATTTAATAAC ACTATTGTTA TGATTACTGA TGTGCATGGT AATGCAATTG CTTGGTCATC | 840 |
| AGCTGGTGCT CTTGGTTTCA AAGGTTCTCG TAAATCTACA CCATTCGCTG CTCAAATGGC | 900 |
| TTCTGAAGCT GCTGCTAAAT CTGCACAAGA ACACGGTCTT AAATCAGTTG AAGTTACTGT | 960 |
| AAAAGGTCCA GGTTCCTGGTC GTGAGTCAGC TATTCGTGCG CTTGCTGCCG CTGGTCTTGA | 1020 |
| AGTAACAGCA ATTCGTGATG TGA CTCCAGT GCCACACAAT GGTGCTCGTC CTCCAAAACG | 1080 |
| TCGCCGTGTA TAATCATCGC ATTACACTGC TTTTCGTTTA AGAGGGAGTA ACTAAATGAT | 1140 |
| CGAGTTTGAA AAACCAAATA TAACAAAAT TGATGAAAAT AAAGATTATG GCAAGTTTGT | 1200 |
| AATCGAACCA CTTGAACGTG GCTACGGTAC AACTCTTGGT AACTCTCTTC GTCGTGTA CT | 1260 |
| TCTAGCTTCT CTACCAGGAG CAGCTGTGAC ATCTATCAAC ATTGATGGTG TGTTACATGA | 1320 |
| GTTTGACACA GTTCCAGGTG TTCGTGAAGA CGTGATGCAA ATCATTCTGA ACATTAAAGG | 1380 |
| AATGTCAGTG AAATCGTACG TTGAAGACGA AAAAATCATC GAACTGGATG TTGAAGGTCC | 1440 |
| TGCTGAAGTA ACAGCTGGTG ACATTTTGAC AGATAGCGAT ATTGAAATTG TAAATCCAGA | 1500 |
| TCATTATCTC TTTACAATCG GTGAAGGTTT TTCTCTAAAA GCGACTATGA CTGTTAACAG | 1560 |
| TGGTCGTGGA TATGTACCTG CTGATGAAAA TAAAAAGGAT AATGCACCAG TTGGAACACT | 1620 |
| TGCTGTAGAT TCTATTTATA CACCAGTTAC AAAAGTCAAC TATCAAGTGG AACCTGCTCG | 1680 |
| TGTAGGTAGC AATGATGGTT TCGACAAATT AACCTTGAA ATCTTGACAA ATGGAACAAT | 1740 |
| TATTCCAGAA GATGCTTTAG GGCTTTCAGC ACGTATTTG ACAGAACATC TTGATTTGTT | 1800 |
| TACAAATCTT ACTGAGATTG CTAAGTCAAC TGAAGTGATG AAAGAAGCTG ATACTGAATC | 1860 |
| TGACGACCGT ATTTTAGATC GTACGATTGA GGAACCTGGAC TTGTCTGTGC GTTCATACAA | 1920 |
| CTGTTTAAAA CGTGCCGGTA TCAATACTGT GCATGATTTG ACAGAAAAAT CTGAAGCAGA | 1980 |
| GATGATGAAA GTACGAAATC TTGGACGCAA GAGTTTGGA GAAGTGAAAC TCAAATCAT | 2040 |
| TGATTTGGGT CTTGGATTAA AAGATAAATA AAGGAGGAAT ACATGGCTTA CCGTAAACTA | 2100 |
| GGACGCACTA GCTCACAACG TAAAGCAATG CTTGCGGATT TGACAACTGA CCTTTTGATC | 2160 |
| AACGAATCAA TCGTGACAAC TGAAGCTCGT GCTAAAGAAA TCCGTAAAAC TGTTGAAAAA | 2220 |

878

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| ATGATTACTC | TAGGTAAACG | TGGTGATTTC | CATGCACGTC | GTCAAGCAGC | TGCTTTCGTA | 2280 |
| CGTAATGAAA | TCGCATCTGA | AAACTATGAT | GAAGCAACTG | ATAAGTACAC | TTCTACTACA | 2340 |
| GCACTTCAAA | AATTGTTCTC | AGAAATCGCA | CCTCGTTATG | CTGAACGTAA | CGGTGGATAC | 2400 |
| ACTCGTATCC | TTAAAACTGA | ATCACGTCGT | GGTGATGCAG | CGCCAATGGC | GATCATCGAA | 2460 |
| TTAGTATAAA | ATCATCAATT | TTGTTGAGTG | TTATGATGAT | GGAGTCTTGT | GCTCTTAGTC | 2520 |
| TAGCTCTGGT | CTACCGCTAG | GATTTCGGTC | CTAGCGGGAA | CACTCATCAT | AAGTTGGGAT | 2580 |
| AGTAGACGCT | TGTTTACGAA | ATTGTTTTTT | TCTTAAGAAC | AACTTCGTAA | GCAGGCGTTT | 2640 |
| TTGAGTATTT | TCGTTAGAAT | TATGCTATAC | TATTTGAAAA | GAATCCTGTT | TAATGTTAAG | 2700 |
| GTTTCTTAT | TTAAGAAGAA | TTGGAGTTTA | CTTATGAAAG | CCATTATAAC | TGTTGTGGT | 2760 |
| AAAGATAAAT | CTGGAATTGT | TGCAGGTGTT | TCTGGTAAAA | TTGCAGAATT | AGGATTGAAT | 2820 |
| ATTGACGATA | TCTCTCAAAC | TGTCTTGGAT | GAATATTTTA | CGATGATGGC | TGTTGTATCT | 2880 |
| AGTGATGAAA | AGCAAGATTT | TACCTATCTT | CGTAATGAAT | TTGAAGCTTT | TGGGCAAACT | 2940 |
| TTGAATGTAA | AAATCAATAT | TCAGAGTGCA | GCGATTTTCG | AAGCTATGTA | TAATATCTAG | 3000 |
| GAGGTCATCA | TGGATATTAG | ACAAGTTACT | GAAACCATCG | CCATGATTGA | GGAGCAAAAC | 3060 |
| TTCGATATTA | GAACCATTAC | CATGGGGATT | TCTCTTTTGG | ACTGTATCGA | TCCAGATATC | 3120 |
| AATCGTGCTG | CGGAGAAAA | CTATCAAAAA | ATTACGACAA | AGGCGGCTAA | TTAGTAGCT | 3180 |
| GTTGGTGATG | AAATGCGGC | TGAGTTGGGA | ATTCTATCG | TTAATAAGCG | TGTATCGGTG | 3240 |
| ACACCTATTT | CTCTGATTGG | GGCAGCGACA | GATGCGACGG | ACTACGTGGT | TCTGGCAAAA | 3300 |
| GCGCTTGATA | AGGCTGCGAA | AGAGATTGGT | GTGGACTTTA | TTGGTGGTTT | TTCTGCCTTA | 3360 |
| GTACAAAAAG | GTATCAAAA | GGGAGATGAG | ATTCTCATCA | ATTCCATTCC | TCGCGCTTTG | 3420 |
| GCTGAGACGG | ATAAGGTCTG | CTCGTCAGTC | AATATCGGCT | CAACCAAGTC | TGGTATTAAT | 3480 |
| ATGACGGCTG | TGGCAGATAT | GGGACGAATT | ATCAAGGAAA | CAGCAAATCT | TTCAGATATG | 3540 |
| GGAGTGGCCA | AGTTGGTTGT | ATTCGCTAAT | GCTGTTGAGG | ACAATCCATT | TATGGCGGGT | 3600 |
| GCCTTTCATG | GTGTTGGGGA | AGCAGATGTT | ATCATCAATG | TCGGAGTTTC | TGGTCTTGGT | 3660 |
| GTTGTGAAAC | GTGCTTTGGA | AAAAGTTCGT | GGACAGAGCT | TTGATGTAGT | AGCCGAAACA | 3720 |
| GTTAAGAAAA | CTGCCTTTAA | AATCACTCGT | ATCGGTCAAT | TGGTTGGTCA | AATGGCCAGT | 3780 |
| GAGAGACTGG | GTGTGGAGTT | TGGTATTGTG | GACTTGAGTT | TGGCACCAAC | CCCTGCGGTT | 3840 |
| GGAGACTCTG | TGGCACGTGT | CCTTGAGGAA | ATGGGGCTAG | AAACAGTTGG | CACGCATGGA | 3900 |
| ACGACGGCTG | CCTTGGCCCT | CTTGAACGAC | CAAGTTAAAA | AGGGTGGAGT | GATGGCCTGC | 3960 |
| AACCAAGTCG | GTGGTTTATC | TGGTGCCTTT | ATCCCTGTTT | CTGAGGATGA | AGGAATGATT | 4020 |

| | |
|--|------|
| GCTGCAGTGC AAAATGGCTC TCTTAATTTA GAAAACTAG AAGCTATGAC GGCTATCTGT | 4080 |
| TCTGTTGGAT TGGATATGAT TGCCATCCCA GAAGATACGC CTGCTGAAAC TATTGCGGCT | 4140 |
| ATGATTGCGG ATGAAGCAGC AATCGGTGTT ATCAACATGA AAACAACAGC TGTTCTGTATC | 4200 |
| ATTCCCAAAG GAAAAGAAGG CGATATGATT GAGTTTGGTG GTCTATTAGG AACTGCACCC | 4260 |
| GTTATGAAGG TTAATGGGGC TTCGTCTGTC GACTTCATCT CTCGCGGTGG ACAAATCCCA | 4320 |
| GCACCAATTC ATAGTTTAA AAATTAAGAA AATAGGAGAA ATTTTAAGTT CTATTTAAGA | 4380 |
| TTAGACGTGT ATACTATAAT CATTAAATAA AGACCTCCTA ATATTATTTG AAACAGATAA | 4440 |
| CACTGAATTA GTTTGAATTT GATTTTCATC TAATATCTTT ATTTAATGAA CTCCTAAACT | 4500 |
| TTTTCATAAT AATCTCCTTC AAAAGTCGCC TGTATGGGTG GCTTTTATTT TATCATTCAT | 4560 |
| GATATAATAG AAGCAAACGG AGGACGGAAA ATGGTAAAG TACGATTGTA TTTGGTACGT | 4620 |
| CATGGCAAGA CCATGTTTAA CACGATTGGT CGCGCGCAAG GTTGGAGCGA TACTCCCTTA | 4680 |
| ACTGCTGAAG GTGAACGAGG GATTCAAGAG TTAGGAATCG GTTTGCGAGA ATCTGATCTA | 4740 |
| CAGTTTGAGC GTGCTTATTC GAGTGATTCT GGTCTACCA TTCAGACCAT GGAATTATC | 4800 |
| CTTGAAGAAC TTGGCTTGCA GGGGGAAATC CCTTATCGCA TGGACAAGCG TATCAGAGAA | 4860 |
| TGGTGTTCG GTAGTTTGA TGGAGCCTAT GATGGCGATC TTTTCATGGG CATTATTCCT | 4920 |
| CGTATCTTTA ATGTGGACCA CGTTCACCAA TTGTCTTATG CTGAACTGGC TGAGGGCTTG | 4980 |
| GTAGAGGTCG ATACAGCTGG TTGGGCTGAA GGCTGGGAAA AACTCAGTGG CCGAATCAAG | 5040 |
| GAAGGCTTTG AAATGATTGC AAAAGAAATG GAAGATCAAG GTGGAGGTAA CGCCCTTGTT | 5100 |
| GTCAGCCATG GAATGACTAT TGAACCATT GTTTATCTGA TTAATGGCAT GCATCCGCAT | 5160 |
| GGTCTGGATA ATGGTAGCGT GACAATCCTT GAATATGAGG ACGGCCAGTT TAGGGTTGAA | 5220 |
| GTTGTCCGTG ACCGTAGTTA CCGAGAGCTA GGACGTGAGA AGATGGAAGA AGGCTCTATT | 5280 |
| TAATCAGTCT AGACTTGCTT GCCATGAGCT AGGGATTTGA TAAGAATATC AAGATAAGAA | 5340 |
| AAAACAGCCG AGGGCACTCC TTTCGGCTGT TTTTGATGTG GAAAACTAAA GTGTAATGCT | 5400 |
| ATTGCTTTTA GAGATTTTCA TAAACAAGAG CAAGGAACCT ACTGTTAGAA CAGTCAGGAT | 5460 |
| AGTTGACAAG GTTGCGGCTA CACCGTAATT TCCTCTGAGA ACCTCTGTAT AAATAGCTAC | 5520 |
| AGTCATTGTT CTTGTTTGA CATTGTAGAG GAGGATAGAA GTAGAGAGTT TTGAAATCAT | 5580 |
| TGTGACTCAA GATAAGATGG CTCCAGAAAT GATACCAGAT AGCATCATTG GAGTTGTAAT | 5640 |
| CTTAGCAAAG GTATTGAGAC GACTACTTCC TAAGCTTTCA GCAGCTTCTT CAATACTTGG | 5700 |
| TGCTATTTGT TGTAAGCTAG CAACAGATGA GCGAATAGTA TAAGGTAATC TTCTGGCAGA | 5760 |

880

| | |
|--|------|
| TAGAGACATA ATCAAGATGA AAGCAGTCCC TGTAAATCATA AGAAATCCAC TTCCAAATAG | 5820 |
| ACCAGTATTG AAGGAAGAAA TGAAGGCAAT CCCTAGAACG GTTCCTGGTA CAATATAAGG | 5880 |
| TACCATACTG AGGCTGTCAA TTAAGTTTGT AAACAAATTC CGTTTTCTAA CGGCTAGGTA | 5940 |
| GGAGATAAAT GTCGCAAATA GAACAAC TAG AACTAAGGCA ATCAAAGGGA TACGAATGGT | 6000 |
| ATTGAAAATA GCAGATCCCA TACGATGGAA AGCTACCTTG TAACTGTTTG GAGAATAACC | 6060 |
| TTTAACAGAT ACCATACCTG ATGTTTTTAG GAAAGAGGTA TAAATTAAGT AGATTTGAGG | 6120 |
| TAAAACAGAG ATAAAGATAA TTCCGTAGAC TGTGTCATAA ATGGCAGCCA TTTTTCCTTT | 6180 |
| TGTAGTTTTT TTAGGCTCAA TTGGATGGAG CAGATTCATG CTGAAACTGT AGCGGTTTGC | 6240 |
| AATGTGTTTT TGGATAAGGA AAATTGCCAA GGCAATGATA ATCGCCATAA TTGCAAAAGC | 6300 |
| AGAATTTCTT CCAACCTCGC TAATAAATTG GGTATAAATC AGGACAGGGA AAGTCCGATA | 6360 |
| CCCTTCGCCA ATCAACATAG GCGTCCAAA GTCGTAGAAT GCTCTCATAA ATACAAGCAA | 6420 |
| GGAGCTGCTA GTAAGGTTGG AACTAGGAGA GGTAACAAC CCGTTACGAT AGGTTTAAAT | 6480 |
| CCGAAGGACC CCATGCTTTC AGCTGCTTCA AGTAGAGAAT TGTCAATACT GTTCATTGTT | 6540 |
| CCAGCAACAT ATAGAAATAC CAGTGGGAAT AGTTGCAGTG TAAAGACAAG TACAATTCTT | 6600 |
| TTGAATCAAT AAATATCGAT AGCTGGAAGA TAAAGGGCAT TTGTCAAAAA TTTAGTGATG | 6660 |
| ACCTCATTTT GCCTAGCAA GAGAAGCCAG GAGTAGGCTC CTACGAAAGG AGCTGACATG | 6720 |
| GAAGCAATGA TAATCAATAT TTGTAGAAAT TTCTTCCCCT TGAAGTCATA CATAGAGAAG | 6780 |
| AGATAAGCTA ATAGGGTTCC TACAAC TAAG GAAGTGATAG TAGCGGTAAT GGAAACCTTG | 6840 |
| AAACTGTTGA CTAGTGCTC AGAGTAGTAG GCTTTACTAA AGAAAGTGAC AAAATTAGCT | 6900 |
| AGTGAGAATT GTCCTTCATG TATAAGTGCT TGCTTGAGCA CGGTAACGAT AGGATAAACG | 6960 |
| AGAAAGATAG GATAGGTAAG AAAGAGGAAG AAAGAGGAAA CTGTCCAAAT ATTTAGTTTT | 7020 |
| TTACGTTCCA TGGTTGACTC CTTTTATCAG GTTTTGGGAA CCATCTGCAG AAAAGATGTT | 7080 |
| TAATTTTTGC GTATTGATTC GTAGACGAAT ACGATTGCCT TTTTGTAGAT CTCTTCAAA | 7140 |
| AGTTGATTCT TCACTAACTT GAATTTTTGA GGCAAAACCT GTCTCAATGA AATAATCCGT | 7200 |
| ATTTAGTCCA AGATAGACGC TATCTCTAAT AGTTCCTTCA ATATCTCCAG ATTCATCTTT | 7260 |
| GATAAACTCT TCGGGACGAA TGCTTACATG AATAGCTTGC TCCTCAACCT GATCAAGAGC | 7320 |
| TGGCATTCTGA AGGGCATAGC CATCTGAAAA GACGATATAA GCGCCGTCGC TCCGTTTTTC | 7380 |
| AAGATTGGCA GGGATAATAT TTGTGCGTCC GATAAAGGTT GCCACAACT CATTAGCTGG | 7440 |
| TTTATGATAG AGTTCTTTTG GTCGGCCGAT TTGTTGGATC ACCCATCTT TCATAACAGC | 7500 |
| AATTTGGTCT GAAATAGCCA TGGCTTCTT TTGTCGTGG GTTACATAAA CAGTTGTAAT | 7560 |

881

| | | | | | | |
|-------------|------------|------------|------------|------------|-------------|------|
| TCCCACCTTCG | TGTTGGATTT | CTCGGATGGC | TTGACGCATA | TCCAAGCGAA | GTTTGGCCTC | 7620 |
| CAGATTACTA | AGTGGCTCGT | CCATGAGGAG | AACACTTGGA | TTAACCGCTA | AGGCCCATGC | 7680 |
| CAAGGTGACA | CGTTGTTGTT | GTCCACCACT | GAGTTTATCG | GGCTTTCGAT | CCGCATATTG | 7740 |
| AGCAATTGTC | ATGAGTTCAA | GATACTTGTT | GGTCTGTTGA | ATCAATTCTT | CTTTTGGAAC | 7800 |
| CTTCTTTTGC | ATAAGACCAA | AAGCAACGTT | GTCTCGGACA | GTCAAATGTG | GGAAAATAGC | 7860 |
| GTAGTTTTGG | AAAACCATCC | CGATATTGCG | TTTGCTGGGT | TCCATATTAT | TGATTTTTGT | 7920 |
| ATCATCGAAG | TAAAATTCTC | CACCTTCGAT | ACTGTTGAAA | CCTGCAATCA | TACGAAGAAG | 7980 |
| GGTCGTTTTC | CCACATCCTG | AAGCTCCAAG | AAGGTAAG | AGACTTCCTT | TTGGAATTGT | 8040 |
| AATGTTCAAA | TTCTCAATAA | CAGGGACATC | GTGGTAGATT | TTTTTGCGT | TAATAATTTT | 8100 |
| GATCTCACTC | ATAGTGAACC | TCTTTTACTG | TTTAGATTGG | ATATCTGTAA | AGACTTCGTT | 8160 |
| GTTTTCTTA | ACGATATCTG | ATTTATTCTT | GATGACATAA | TCATAATCTT | CAGTGAGTGT | 8220 |
| TTTGATTTTG | TCAATTGGTT | TCATGTTTTC | GCTTGTTTTA | GCATTTTTAC | GAACAGGACG | 8280 |
| GTTAGTAGTG | GTTGTACCAA | GTGTATCTTG | TACTTCTTGA | GAGATAATAA | AATCGATAAA | 8340 |
| TTTCTTGGA | TTTTCCATAT | TTTTAGATTT | TTTAACGATA | GCAGCACTAG | CAGGTAGGAA | 8400 |
| GACGGTTCCT | TCTTTTGAT | AGACTACCTT | AATGTTAGCT | CCGTCATTTA | AGAGTTTAAAC | 8460 |
| TGCTGGATCT | TCATAAGAGA | GACCAACAGC | CATTCTCCA | TCAGCGACTA | CTTTATAGAC | 8520 |
| ACTAGATGAA | CTTGAACCGA | TTTTACCATC | AATAAGTGTG | AAAAGATCTT | TTACATAAGA | 8580 |
| CCAAGCCTTA | TCATCTTTGT | AACCACCTTG | AGCTGTAGC | ATATTTGTTA | ATTGAGCAAA | 8640 |
| GGCGCTAGAA | GAGTTTGCTG | GGTCAGCAGT | TGCGATTTTT | CCTTTTAGTT | CAGGTTTGAA | 8700 |
| AAGATCGTTA | TATCCTTCGA | TGTTCATGCC | TTTAGTTAAA | TCAGGGTTGA | CGATTAAAAC | 8760 |
| ACTACCATCT | AGTGATAAG | GAGTAGAGTA | GCCAGTTGTG | TTTTGATATT | CTTTGATAAC | 8820 |
| ATTATCATTT | TCTTTTGAAG | TATAGTTTTC | AAAGAGTTCT | CCGTGGGTAG | TATATTGTGT | 8880 |
| ATAAGAACCA | CCAAAGATAA | CATCAGCTAC | AGGAACTTCT | TTTTCTGACT | CTAGTTTTTT | 8940 |
| GAAAAGTTCT | CCAGTACCAG | CTTGAATCAG | TTCTACTTTG | ATACCATATT | TTTCTTCAAA | 9000 |
| GGCAGGAATA | GTTGCTCCAA | TTAAGCCCTC | TGAGTTTGGT | GAATAAACGA | CTAGCGAACC | 9060 |
| GGCGTCTCCT | TTATCAGATG | AACTGTCATC | GGCAGATTCA | TTAGAAGAAC | AAGCAGCATA | 9120 |
| ATACATCCAT | TTCTTTTTCA | TGATGGATAC | CTCCGTTGTG | TTATTTAAGT | TTATTTTAAA | 9180 |
| ACAATGTAAG | CGTTTTTAAA | ACATACAATT | CTATTCTATA | GTGTATTGAA | TCTATAACAG | 9240 |
| TACACTTTGA | CTGCTAAAAT | ATTTCTATAA | ATTAATTTGA | CTTTCCTGAT | AGAGATGTTC | 9300 |

882

| | | | |
|-----------------------|-----------------------|-----------------------|------|
| ACATCTTATT TCAATTCAC | ATATTAGAGT AAAATTCTCT | ACAAAAAGAA GAATAGCCTA | 9360 |
| TTTTACTATT CTTCTGAGT | ATTTCAAATC CTTGGGGAA | ATATGGAGAT ACTTTTTAAA | 9420 |
| TCCTGACAAA TGGTGTGTT | CTTTTCTAAA TCGGTGATAC | TGTATCGGAG AATGCGCGTG | 9480 |
| AGGTCACAAA GGCTGCGATA | GAGCTTCTAT GGAGAATTTC | TTTTTGGAGA GATTTTTTAA | 9540 |
| AGGAATGAGA CATCCGCTAC | CTCCTTGGA GGTTTTTG | | 9578 |

(2) INFORMATION FOR SEQ ID NO: 128:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 13440 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 128:

| | | | |
|-----------------------|-----------------------|-----------------------|------|
| CGGGCTGTTG TGACGATTCT | TATTTCTATC TGTGTTATCT | TTTTGGGAAC TATTTTGGGT | 60 |
| GTGTGCTTGG CTTTGGGCA | ACGTTCAAAG TTTAAACCGC | TGTTTGGTT GGCCAACTTG | 120 |
| TACGTTTGGG TTTCCGTGG | GACACCGATG ATGGTTCAAA | TTATGATTGC CTTGCTCTT | 180 |
| ATGCATATCA ATGCTCCGAC | TATTCAGATT GGAATTTTAG | GTGTTGATTT TTCGCGTCTG | 240 |
| ATTCCAGGGA TTTGATTAT | CTCTATGAAT AGTGGTGCTT | ATGTTTCGGA GACTGTTCTG | 300 |
| GCCGGAATCA ATGCGGTTC | AAAAGGTCAG CTAGAAGCGG | CTTATTCGCT AGGGATTCTG | 360 |
| CCTAAAAATG CGATGCGTTA | TGTGATTTTG CCACAAGCAG | TCAAAAATAT CTTGCCAGCA | 420 |
| TTGGGGAACG AATTATCAC | CATTATCAAG GACAGCTCCC | TCTTATCAGC TATTGGGGTC | 480 |
| ATGGAGTTGT GGAATGGGGC | TACAACAGTT TCTACAACAA | CCTATCTACC TTAAACACCA | 540 |
| CTTTTATTTG CAGCATTTTA | CTACTTGATT ATGACCTCTA | TTCTGACAGT AGCCTTGAAA | 600 |
| GCTTTTGAAA AACATATGGG | ACAAGGAGAT AAGAAATAAT | GACAGAAACC TTGATAAAAA | 660 |
| TTGAAAATTT ACATAAATCC | TTTGGAAGA ATGAAGTATT | GAAGGGCATC AACCTCGAGA | 720 |
| TTAAAGAGG AGAAGTTGTC | GTTATCATCG GTCCTTCAGG | GAGCGGGAAT TCTACCTTGC | 780 |
| TTCGCTCTAT GAATTTGTTG | GAAGAAGCAA CCAAGGGGAA | GGTTATCTTT GAGGGAGTCG | 840 |
| ATATTACGGA CAAGAAGAAT | GACCTGTTTG CCATGCGTGA | GAAGATGGGC ATGGTTTTTC | 900 |
| AACAATTCAT TCTCTTCCT | AATATGACTG TGATGGAAAA | TATCACCTTG TCCCCTATCA | 960 |
| AGACCAAAGG TGACAGTAAG | GCCGTTGCAG AGAAAAGAGC | TCAGGAACCT TTGGAAAAAG | 1020 |
| TTGGTTTGCC AGATAAGGCA | GACGCTTATC CACAGAGTTT | GTCAGGTGGC CAGCAACAGC | 1080 |
| GGATTGCCAT CGCGCGTGGG | TTGGCTATGG AACCAGATGT | TTTGCTCTTT GACGAGCCAA | 1140 |

883

| | |
|--|------|
| CTTCAGCCCT AGATCCTGAG ATGGTTGGAG AAGTTCTGGC TGTTATGCAA GATCTAGCCA | 1200 |
| AGTCAGGAAT GACCATGGTT ATCGTAACAC ATGAGATGGG ATTTGCCCGT GAGGTGGCAG | 1260 |
| ATCGTGTCTAT CTTTATGGCA GACGGTGTGG TTGTTGAAGA CGGAACACCT GAGCAGATTT | 1320 |
| TTGAACAAAC CCAAGGACAA AGGACTAAAG ACTTCTTGAG TAAGGTTTTA TAAGTTAGCT | 1380 |
| TTGTTTAGCT ATTTGTAGCC AGCTTTAAAC GTTAAAGAGA AGATTAGTGA AAAGCTCAAC | 1440 |
| CAGAGCTTTT TCTTATAGTT TAAAGCTATA GGATTGCCTA GGAAAGAAGT GTTAGAGCTA | 1500 |
| CATTGTATTT TTTGGTATAA TTAAAGATAT TTGTAAGAAA AGAGAAGTGA TATGACACAG | 1560 |
| ATTATTGATG GGAAAGCTTT AGCGGCCAAA TTGCAGGGGC AGTTGGCTGA AAAGACTGCA | 1620 |
| AAATTAAAGG AAGAAACAGG TCTAGTGCCT GGTTTGGTAG TGATTTTGGT TGGGGACAAT | 1680 |
| CCAGCCAGCC AAGTCTACGT TCGCAACAAG GAGAGGTCAG CCCTTGCGGC TGGTTTCCGT | 1740 |
| AGCGAAGTAG TACGGGTTC AGAGACCATT ACTCAAGAGG AATTGTTAGA CCTGATTGCT | 1800 |
| AAATACAATC AGGATCCAGC TTGGCATGGG ATTTTGGTTC AGTTGCCATT ACCAAAACAC | 1860 |
| ATTGATGAAG AGGCGGTCTT ATTGGCTATT GACCCAGAAA AGGATGTGGA TGGTTTCCAT | 1920 |
| CCTCTAAACA TGGGGCGTCT TTGGTCTGGT CATCCAGTCA TGATTCCTTC GACACCGGCA | 1980 |
| GGAATTATGG AAATGTTCCA TGAATATGGG ATTGACTTGG AAGGTAAAAA TGCAGTCGTC | 2040 |
| ATCGGTGCGAT CCAATATTGT CGGAAAACCT ATGGCCAGC TTCTTTTGGC AAAGAATGCA | 2100 |
| ACAGTAACCT TGACTCACTC ACGTACTCAT AATCTTTCCA AGGTGGCTGC AAAAGCAGAT | 2160 |
| ATTCTGGTTG TTGCAATCGG TCGTGCCAAG TTTGTGACTG CTGACTTTGT CAAACCAGGT | 2220 |
| GCGGTAGTCA TTGACGTTGG GATGAACCGC GATGAAAATG GTAAGCTCTG TGGGGATGTT | 2280 |
| GATTATGAGG CGGTTGCCCC ACTTGCTAGC CACATTACGC CAGTCCCTGG AGGTGTCGGT | 2340 |
| CCTATGACCA TTACTATGCT GATGGAGCAA ACCTATCAGG CAGCACTTAG GACATTGGAT | 2400 |
| AGAAAATAAG ATAAAAATTT TCTGAGGAAA GTGTATTTTC TATAGCTATA TCTAAAATGA | 2460 |
| TAGAAATGAA TATTAAATTT TAGAAATAAG TTTATAAAAG GAGGTTTGCG CCTCCTTTTT | 2520 |
| GTTGTATAAT GGAGTGAGGT GATTAGATGA TTTTAAAAAT TTATAATGGG GAATATAGTT | 2580 |
| TACAAATGGGA TGGAATATAC TACTTAGCAC TAATTGATTA TCCAAATATT CAAGAGTGGG | 2640 |
| AATTAGAAAA AATTGCTAAA TTTATAGCTT ACGAAAACT TCATAAACGT CAAACAAGTA | 2700 |
| TTGAGTGTGC TGATTCTTGT TTAAAAAAG AAATTTTAGA TTACATCTGT CAGCATCCCT | 2760 |
| TTCTGCCACC ATTTACTCCT ACAGATAAAA GAGTAGCCTC GACTTATGAC CTACATAAGA | 2820 |
| GGTTAGTGAC TTCAGACTAC TGTAGTCATA CTACGACTAT AGATGCAGCG ATTTCTATTT | 2880 |

884

| | |
|--|------|
| TTAAACTGG TCGTCTTTTA TCTGCTGTGA AAGCCTTTGG GCGAGATGCT GAGGAGTTGG | 2940 |
| TTTTGGATAG TCGAAATGCT GCATCTGATC CGATAGATTA TTTTGACTAT GTCATGTTAG | 3000 |
| GGTGGTCAAA TACAAGTTCT GGTATCGAT TGGCGATGGA GCGTTTATTA GGTCGAGCTC | 3060 |
| CTTCAGAGAA AGAATTACAA GACAAGTTTA TTCCTGGAGT AAGTTTTCAT TTTATCTATA | 3120 |
| CAGATTTGAT TAAAGTTCCT GGTATATTTT TTGATGGTTA CCATGCTGTA AAAATTAAGG | 3180 |
| ACATGCTTAA TTTATTAAGT GAGTTGTATA TTTGCATTAT TCCAACATCAT AATAAGAGCC | 3240 |
| AATTTGAAAA TATTATTCCA ACCAAAATAC AAGATAGGGT GTATTATCTT GACTATGCTG | 3300 |
| GAGAAGACTT AGAAGAGTGG ACTAAGAAAG TCTATCAAGT TGTTTTAAAA CAATCAGATA | 3360 |
| AAGGATAGTT GAGGAAAAAA CGATGAAAGT GATTGATCAA ACCTTACTAG AAAAAGTCAT | 3420 |
| TATTGAACGT TCTTGTACAA GTCATAAAGG AGACTACGGT CGTCTGCTGT TGCTTGGTGG | 3480 |
| GACTTATCCT TATGGTGGTG CCATCATCAT GGCTGCTTTA GCAGCTGTAA AAAGCGGTGC | 3540 |
| AGGATTGGTA ACCGTGGGAA CGGACAGGGA AAATATCCCT GCTCTACACA GCCATTTGCC | 3600 |
| TGAGGCTATG GCCTTTTCTC TGCAAGATCA GTAATTGTGA CAAGAGCAAT TGGAGAAGGC | 3660 |
| AGAAGTTGTC TTGCTGGGGC CTGGTTTACG AGACGATACG TTTGGAGAAA ATCTTGTA | 3720 |
| ACAGGTCTTT GCTAGCTTAA AAAAGAATCA GATTTTGATT GTAGATGGAG GGGCCTTAAC | 3780 |
| CATCCTTGCT AGGACAAGTT TGTGTGTTCC ATCTAACCAG CTTATCTTAA CTCCTCCCA | 3840 |
| AAAAGAATGG GAAAACTGT CTGGTATGTC TATTGAAAAG CAAAACGAAG GTACAACATC | 3900 |
| TAGTGCCCTG ACTTCTTTCC CTCAAGGAAC AATTTTGGTA GAGAAAGGTC CAGCTACTCG | 3960 |
| TATTTGGCAA GTTGGCCAGT CTGATTATTA CCAGTTAAAG GTTGGCGGTC CCTATCAGGC | 4020 |
| GACTGGTGGT ATGGGTGATA CACTGGCTGG AATGATTGCA GGATTTGCAG GCCAATTTG | 4080 |
| ACAGGCCAGT CTCTACGAAC GTGTGGCAGT AGCAACCCAT CTTCAATTCAG CCATAGCCCA | 4140 |
| AGAACTATCT CAAGAAAAT ATGTGGTCTT GCCGACGGAA ATTAGTAATT GTCTTCTTAA | 4200 |
| AGTAATGAAA AGATATGTCT AAAATAGTTA GACAAAAAAT GTTGATAATT TGTATCATTA | 4260 |
| TTCTTAATTC ACAAAAAACG AACGTTTAGT ATTCTTCTTG CTAAGAAACT AAATTTGTTC | 4320 |
| GTTTTTTTAC TCTTGTAAT CTATTTTGT TAGAGTTGAT TTGGTTTACA TCCGTACTTA | 4380 |
| AATTGATTG TTAGAGCTCT ACTTTTATTA AAAAAATTCA ATTTCAAGGA TAAATAAGCA | 4440 |
| GTATTCTAAA GGTACTTTTA GATGAAATAA AAGCCTTTAC ATGGTATAAT AGAGGTAGCT | 4500 |
| CTTTAATGGA GGTGTTTGAG TGGAAAATCT GAAGAAAATG GCAGGTATCA CGGCTGCTGA | 4560 |
| ATTTATCAAG GATGGGATGG TTGTAGGCT AGGAACAGGT TCTACTGCCT ATTATTTTGT | 4620 |
| CGAAGAAATC GGTGTCGAA TCAAGGAAGA AGGCTTGCAG ATTACAGCTG TGACGACTTC | 4680 |

885

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| TAGTGTGACC | AGTAAACAGG | CTGAAGGGCT | CAATATCCCG | CTCAAGTCTA | TTGACCAAGT | 4740 |
| AGACTTTGTC | GATGTGACAG | TCGACGGGGC | GGATGAAGTG | GATAGTCAGT | TTAATGGAAT | 4800 |
| CAAAGGCGGT | GGTGGTGCCC | TTCTCATGGA | AAAGGTGGTC | GCAACACCAT | CAAAAGAATA | 4860 |
| CATTTGGGTG | GTGGATGAAA | GCAAGCTGGT | CGAAAACTA | GGTGCTTTTA | AATTGCCAGT | 4920 |
| AGAAGTGGTT | CAGTATGGTG | CAGAGCAGGT | CTTTCGTGAT | TTTGAACGAG | CTGGCTACAA | 4980 |
| ACCAAGTTTC | CGTGAAAAAG | ACGGCCAACG | TTTTGTGACC | GATATGCAGA | ATTTTATCAT | 5040 |
| TGACCTCGCC | TTGGATGTCA | TTGAAAATCC | AATTGCTTTT | GGACAAGAAT | TGGACCATGT | 5100 |
| CGTTGGTGTT | GTGGAGCATG | GTTTATTCAA | CCAAATGGTG | GATAAGGTAA | TCGTTGCTGG | 5160 |
| ACGAGATGGA | GTTTCAGATT | CAACTTCAAA | AAAAGGAAAA | TAGAAGGGGG | CATAAGATGT | 5220 |
| CTAAATTTAA | TCGTATTGAT | TTGGTGGTAC | TGGATTCTGT | AGGAATCGGT | GCAGCACCAG | 5280 |
| ATGCTAATAA | CTTTGTCAAT | GCAGGGGTTT | CAGATGGAGC | TTCTGACACA | CTGGGACACA | 5340 |
| TTTCAAAAAC | AGTTGGTTTG | AATGTCCCAA | ACATGGCTAA | AATAGGTCTT | GGAAATATTC | 5400 |
| CTCGTGAAAC | TCCTCTTAAG | ACTGTAGCAG | CTGAAAGCAA | TCCAACTGGA | TATGCAACAA | 5460 |
| AATTAGAGGA | AGTATCTCTT | GGTAAGGATA | CTATGACTGG | ACACTGGGAA | ATCATGGGAC | 5520 |
| TCAACATTAC | TGAGCCTTTC | GATACTTTCT | GGAAACGGAT | CCCAGAAGAA | ATCCTGACAA | 5580 |
| AAATCGAAGA | ATTCTCAGGA | CGCAAGGTTA | TTCTGTAAGC | CAACAAACCT | TATTCAGGAA | 5640 |
| CGGCTGTTAT | CTATGATTTT | GGACCACGTC | AGATGGAAAC | TGGAGAGTTG | ATTATCTATA | 5700 |
| CTTCAGCTGA | CCCTGTTTTG | CAGATTGCTG | CCCACGAAGA | CATTATTCTT | TTGGATGAAT | 5760 |
| TGTACCGTAT | CTGTGAATAC | GCTCGTTCGA | TTACCCCTGA | GCGTCCTGCC | CTTCTTGATC | 5820 |
| GCATCATTGC | TCGCCCTTAT | GTAGGTGAAC | CAGGTAACTT | CACTCGTACG | GCAAACCGTC | 5880 |
| GTGACTTGGC | TGTATCTCCA | TTTTTCCCAA | CTGTTTTGGA | TAAATTGAAT | GAGGCTGGTA | 5940 |
| TCGATACTTA | TGCTGTGGGT | AAAATCAACG | ATATCTTTAA | CGGTGCTGGT | ATCAACCATG | 6000 |
| ACATGGGTCA | CAACAAGTCA | AATAGTCATG | GAATTGATAC | ACTATTGAAG | ACTATGGGAC | 6060 |
| TTGCTGAGTT | TGAAAAAGGA | TTCTCATTCA | CAACCTAGT | TGACTTTGAT | GCCCTTTACG | 6120 |
| GCCATCGTCG | TAATGCTCAC | GGTTACCGTG | ATTGCTTGCA | TGAGTTTGAT | GAACGCTTAC | 6180 |
| CTGAAATTAT | CGCAGCTATG | AGAGAGAATG | ACCTCTCTTT | GATTACTGCG | GACCATGGAA | 6240 |
| ATGACCCAAC | GTATGCAGGA | ACGGATCACA | CTCGGGAATA | TATTCCATTG | TTGGCCTATA | 6300 |
| GCCCTGCCTT | TAAAGGAAAT | GGTCTCATTC | CAGTAGGACA | TTTTGCAGAT | ATTTACGCGA | 6360 |
| CTGTTGCCGA | TAACTTTGGT | GTGGAAACTG | CTATGATTGG | GGAAAGTTTC | TTAGATAAAT | 6420 |

886

| | | | | | | |
|--------------|------------|-------------|------------|------------|------------|------|
| TGGTATAAGA | TGACGCGCTA | TGCTTTGCTG | GTGAGAGGTA | TCAATGTTGG | TGGTAAGAAT | 6480 |
| AAGGTCGTCA | TGGCGGAGCT | TCGTCAAGAA | TTGACAAACT | TGGGACTGGA | AAAGGTTGAG | 6540 |
| AGCTACATCA | ATAGTGGCAA | TATTTTCTTT | ACTTCGATAG | ATTCCAAAGC | CCAATTGGTT | 6600 |
| GAAAAGCTAG | AGACTTTCTT | TGCAGTCCAT | TATCCATTTA | TTCAGAGCTT | TTCTTTACTG | 6660 |
| AGTCTAGAGG | ACTTTGAGGC | GGAAGTTGAA | AATCTACCAG | CTTGGTGGAG | CAGAGACTTG | 6720 |
| GCACGAAAAG | ATTTTCTCTT | TTACACTGAG | GGTTTGGATG | TGGACCAAGT | CATCGCGACA | 6780 |
| GTTGAAAAGT | TAGAGCTGAA | AGATGAAGTG | CTTTATTTTG | GAAAAGTTGG | GATTTTCTGG | 6840 |
| GGGAAATTTT | CTGAAGAATC | CTATTCTAAG | ACTGCCTATC | ATAAGTACTT | GCTGAAGGTG | 6900 |
| CCTTTCTACC | GCCACATTAC | TATTTCGTAAT | GCTAAAACCT | TTGACAAAAT | TGGTCAAATG | 6960 |
| CTAAAAAAT | AATAAAGGAG | ACACACAATG | ACATTTTAA | ACAAAATCCA | TGAAAGTCT | 7020 |
| ACTTTCCTGA | AAGAAAAGGG | AATGCAGCC | CCTGAGTTCG | GTCTAATCCT | TGGATCAGGA | 7080 |
| CTTGGAGAAT | TGGCAGAAGA | AATCGAAAAT | CCAGTTGTAG | TAGACTATGC | TGAGATTCCA | 7140 |
| AAC TG GGGCC | GTTCAACAGT | AGTCGGTCAT | GCTGGTAAAT | TGGTATATGG | TGAAGTGGCA | 7200 |
| GGTCGCAAGG | TCTTGGCTCT | TCAAGGGCGT | TTCCATTTCT | ATGAAGGGAA | TCCTCTGGAA | 7260 |
| GTGGTGACTT | TCCCAGTTCG | TGTGATGAAA | GTTCTTGGAT | GTGAAGGTGT | TATTGTAAAC | 7320 |
| AATGCAGCTG | GCGGTATCGG | ATTTGGTCCT | GGTACCTTGA | TGGCTATCTC | AGACCATATC | 7380 |
| AACATGACGG | GGCAAAATCC | ATTGATGGGT | GAAAAGTTGG | ATGACTTTGG | CCCACGTTTC | 7440 |
| CCAGATATGT | CTAGGGCCTA | CACACCAGAA | TACCGTGCCA | CTGCCCATGA | AGTGGCTAAA | 7500 |
| AAACTTAATA | TCAAGCTTGA | TGAAGGTGTC | TATATCGGAG | TTACTGGTCC | GACTTATGAA | 7560 |
| ACACCAGCAG | AAATTCGTTT | CTATAAGACA | CTGGGAGCAG | ATGCAGTTGG | TATGTCTACG | 7620 |
| GTTCTCTGAAG | TTATCGTGGC | AGCCCACTCT | GGCTTGAAAG | TTCTGGGAAT | TTCATGTATC | 7680 |
| ACTAAGTTTG | CGGCCGTTT | CCAAGAAGAA | CTCAATCACG | AAGAAGTTGT | AGAAGTGACT | 7740 |
| GAACGTGTTA | AAGGTGATTT | CAAAGGCTTG | CTTAAAGCGA | TTCTTGCTGA | ATTGTAAGAA | 7800 |
| AAAAGATTTA | AAAGGGGGAG | TGCCTCTGTT | TTTTCAGGAT | TGACTGCCTA | TCCGGATTAA | 7860 |
| AGAAGAAACA | GAGGAATACT | ATGAGCTTCT | TCCTGCTCTT | ATAACTGAAA | GAAGCGGAAG | 7920 |
| AATAGGTATG | TCTGATCTGA | TAGCCAGCAT | TGTGAAAGAC | AAGATTCTAG | GATACTAGCA | 7980 |
| TTAGCTTCCT | AGCCAAGCAG | ACTAGTATGA | TAAGGAGAGA | TGAGAATGAA | TTGACTTTCT | 8040 |
| GAATTTCTCA | GTCTTATCAT | ATATAGCACA | ATGAGATTTT | GCTTGAGTCT | GCTTGTAAT | 8100 |
| AAACGAAAAG | AAAGATAAGA | AATAATGAAA | ATTGGTCAAC | GAATTATGCG | CTTTGGCATA | 8160 |
| AAAAATTAAG | TATCGGAGTT | GTATCTGTTG | TAGTCGGCTT | TGATTTCTAG | CTCCAGCTGG | 8220 |

887

| | |
|---|------|
| AATTTTCAGCC AATGAAGTAA AGCAAGATGT AACATCTGAA GTGGTAATAG GTGTGCTAGA | 8280 |
| TTCTAAGGAG GAATTGAAAAG AGTCAGAAAA TGATGCTCCA AAAC TAGAAA CTCCTCTTAG | 8340 |
| AGAGGAGCCA AGACTAGCTC CTCAAACGCT TCCGGAAGCA AGTGAAGTTC TTGAAAACAA | 8400 |
| AAGGGAAGAG TCAAAAGTAG AGATAACATA ACCAGCTCAA GCGGATGATA TCCGCAAGGT | 8460 |
| TGTTGGGGAA TTAGCCAAGG ATATAAGTAT TACTAAGTTG TATATGACAG GTCATTCTCT | 8520 |
| TGGATGTTAC CTAGCTCAGA TTGCAGCGGT TGAAGCTTAC CAAAAATATC CTGATTTTFA | 8580 |
| TAACCATGTA TTGAGGAAAG TGACAACTTT CAGTGCTCCT AAAGTGATTA CTTCCAGAAC | 8640 |
| TGTTTGGAAT GCTAAGAATG GTTCTGGGA TGTGGTTTG GAAAGTCGTA AATTAGCTGT | 8700 |
| TAGTGGAATA ATTAAGCATT ATGTGGTTGA TAATGACAAT GTTGTGACTC CCTTGATTCA | 8760 |
| TAATAATCGT GATATTGTTA CATTTACAGG TAATTCACGC TTAAACACC GTTCTCGTGG | 8820 |
| CTATTTTGAA AGTCCAATGA ATGATATTCC TAACTTAAT ATTGGTAAAC AAGCTACCTT | 8880 |
| GGATAAACAT GGTATATCGT ATCCGAAATT GGATAAAGTG CGATTCTTTA AGAAACAGGC | 8940 |
| TCTGCCTCGA TCTTCTAGTC AACCAAGCGC TGAACCAATG GAAATATTTG CCTCAGGAAA | 9000 |
| ACAGGTTACT CAAAGTTCGA CAGCTTTCGG AGGAGATGCT AGAAGAGCTG TGGATGGCAA | 9060 |
| AGTCGATGGT AACTATGGTC ACAATTCTGT CACTCATACA AACTTCCAAT CTAAGCCTTG | 9120 |
| GTGGCAAGTA GATTTGGCTA AAGAAGAAAC CATTCGCCAA ATCAATATTT ACAACCGAAC | 9180 |
| AGACACTGCC CAGGATAGAT TGGCAAACCT TGATGTCATT CTTT TAGACA GTTCTGGTAA | 9240 |
| AGAAATTGAG TGAAAACGTA TAACATCTCC TAAAGATGTG TCAGCACAAA TTACGATTAA | 9300 |
| CCATAAAAAA GCGCGCTATG TTCGGATTGA GCTAGAAGGC TATAATGCCC TCAGTCTTGC | 9360 |
| AGAAGTTGAA GTTTTCTGCT TTATAGCTAC GAATGCTGAA ACGGCGACAC AAGTTTCTAA | 9420 |
| GCCAGTTCAA CCAATCAGTC AGACTCCTGT GAAGGATAAA ACATTGACAA TTCAACACAG | 9480 |
| TGGAGCTTAC ATTGCCCCT ACTCCATAAC TTGGGAAGAA GTTCCAGTAG ATAAAGATGG | 9540 |
| AAACCAAGTT GTTCGTAGTC ATTCTTGGGA AGGAAGCGGT CGCAACCAGA CTGCAGGTTT | 9600 |
| TGTCTCAAC CTCCCAATCA AAGAAAATAT GAGAAATCTG CGAGTTAAGA TTGAGAAAAA | 9660 |
| GACGGGCCTA CTATGGAATA GATGGCAAAC AATCTATGAA AACAGACCAA TTTTAGCTCA | 9720 |
| ACCCACCGT AAAATTACCC ATTGGGGTAC GACATTGAAT TCCAAGGTGA GTGACGATGA | 9780 |
| TGTCTTGTA TCTGATGGTA GAATGACAGT TAGTTTGTCT AGTTTATAAG AAAGTACTAC | 9840 |
| CTGAGCTTGA ATAGGACTCA GGTAGCTCTC TATGAAAGAA CAAAATTAAT ACTCAATGAA | 9900 |
| AATCAAAGAG CAAACTAAGA AACTAGCCGC AGGTTGCTCA AAGCACTGCT TTGAGGTTGT | 9960 |

888

| | | | | | | |
|-------------|-------------|------------|-------------|------------|-------------|-------|
| AGATAAGACT | GACGAAGTCA | GTCACATATA | TAATCCAAGG | CGACGTTGAC | GTGGTTTGAA | 10020 |
| GAGATTTTCG | AAGAGTATAA | ACAGAAAGGT | AGAGCGCGTG | TTCTAATTG | AACACGAGTA | 10080 |
| GAAAACTTTT | CTAAAAACAA | AAACGAAAGG | ATGGGTAAAC | TGTATTCGCT | GAAGTGAATA | 10140 |
| CGGGCGACTC | TCCTCTAAAT | CAAAATTAAG | AAAGGAATTG | ACCCCACCCT | AAAAGTAGTG | 10200 |
| GGAAAAAGAT | AGTTGATCTA | GCGAGCATCG | CTCACTGCGC | CCAACTCCTA | TTTTCCCTTC | 10260 |
| GCTTTTTGAT | GGGTTTGGA | TCTTTCTCAA | TATAAAATAT | AAAATAAAGA | AAGGTAGAGC | 10320 |
| GTGTGTTTTG | ATTTGAACAC | GAGCGGAAAA | CTCGGAAAAAT | AGATAATCTG | ACTGAAAAAT | 10380 |
| CAGGATTTCT | CGTCAGGTTT | CTAATTTTCA | GTCGTTTTCT | TCTCGCTCTT | TGTATCATAA | 10440 |
| ATTATGTCTA | TCCATATTGC | TGCTCAGCAG | GGTGAATTTG | CTGATAAAAT | TCTTCTTCCT | 10500 |
| GGGGATCCTC | TTCGTGCTAA | GTTTATTGCG | GAGAATTTCC | TTGATGATGC | TGTTTGTTTT | 10560 |
| AACGAAGTGC | GTAACATGTT | TGGTTACACT | GGTACTTACA | AGGGTCACTG | TGTATCTGTC | 10620 |
| ATGGGAACATG | GGATGGGAAT | GCCATCTATT | TCGATTTATG | CGCGTGAGTT | AATCGTAGAC | 10680 |
| TACGGTGTGA | AGAAATTGAT | TCGTGTGGGA | ACTGCAGGTT | CTTTGAATGA | AGAGGTTCAT | 10740 |
| GTTCGTGAAT | TAGTTTGGC | GCAGGCGGCT | GCAACCAACT | CAAACATCGT | TCGTAATGAC | 10800 |
| TGGCCACAGT | ACGATTTTCC | ACAAATTGCT | AGCTTTGATT | TGCTTGATAA | AGCCTACCAT | 10860 |
| ATCGCCAAAA | AACTTGGTAT | GACTIONCAC | GTTGGGAACG | TTTTGTGATC | TGATGCTTTT | 10920 |
| TACTCAAATT | ACTTTGAAAA | GAATATCGAG | CTTGGTAAAT | GGGGAGTCAA | GGCTGTGGAA | 10980 |
| ATGGAAGCAG | CAGCTCTTTA | CTATCTTGCT | GCCCAATACC | ATGTTGATGC | GCTAGCTATC | 11040 |
| ATGACCATCT | CTGATAGCTT | GGTCAATCCA | GACGAAGACA | CAACTGCAGA | AGAACGTCAA | 11100 |
| AATACCTTCA | CTGATATGAT | GAAGGTGGT | TTGGAAACCT | TGATTGCAGA | ATAATTATAG | 11160 |
| CCAAAAAGGG | GCTCTTTGTC | AACTGTAGTG | GGTTGAAAAA | AAGCTAAGCT | TGAGAAAGGA | 11220 |
| CAAATTTTCGT | CCTTTCTTTT | TTGATATTCA | GGGCGATAAA | AATCCGTTTT | TTGAAGTTTT | 11280 |
| CAAAGTTCCG | AAAACCAAAG | GCATTGCGCT | TGATAAGTTT | GATGAGATTA | TTGGTCGCTT | 11340 |
| CCAGTTTGGC | ATTAGAATAG | TGTAGTTGAA | GGGCGTTGAC | GATTTTCTCT | TTGTTCTTTA | 11400 |
| GAAAGGTTTT | AAAGACAGTC | TGAAAAAGAG | GATGAACCTG | CTTCAGATTG | TCCTCAATGA | 11460 |
| GTCCGAAAAA | TTTCTCAGGG | TCTTTGTTCT | GAAAGTGAAA | AAGTAAGAGT | TGATAGATCT | 11520 |
| GATAGTGGTG | TTTCAAGTCT | TCTGAATAGC | TTAAATCTTT | GTCAAGAATT | TCTTTATTTG | 11580 |
| TTAAGTGCAT | GCGAAAAAGTA | GGGCGATAAA | AACGTTTATC | GCTsArTTTA | CGACTATCCT | 11640 |
| GTGGATGAG | TTTCCAGTAA | CGCTTGATAG | CCTTGATATC | ATGAGATTTT | CGTTCAAACCT | 11700 |
| GATTCATAAT | TTGAACACGA | AAACGACTCA | TGGCACGGCT | GAGATGTTGG | ATAATATGGA | 11760 |

889

| | |
|--|-------|
| AACGATCTAG AACGATTTTA GCACACGGAA AAAGCTGTTT AGCCAAGTCA TAGTAAGGAC | 11820 |
| TAAACATATC CATCGTAATG ATTTTCACTT GACAACGAAC GGCTCTATCG TAGCGAAGAA | 11880 |
| AGTGATTTTCG GATGACAGCT TGTGTTCTGC CTTCAAGAAC AGTGATAATA TTAAGATTAT | 11940 |
| CAAAATCTTG CGCAATGAAA CTCATCTTTC CCTTAGTGAA GGCATACTCA TCCCAAGACA | 12000 |
| TAATCTTTGG AAGCCGAGAA AAATCATGCT CAAAGTGAAA GTCATTGAGC TTGCGAATGA | 12060 |
| CAGTTGAAGT TGAAATGGCC AGCTGATGGG CAATATCAGT CATAGAAATT TTTTCAATTA | 12120 |
| ACTTTTGAGC AATTTTTTGG TTGATGATAC GAGGGATTGG GTGATTTTTT TTTACCAGGG | 12180 |
| GAGTCTCAGC AACCATCATT TTTGAASAGT GATAGCACTT GAAACGGCGT TTTCTAAGGA | 12240 |
| GAATTCCTAGA AGGCATACCA GTTGTTCGA GGTAAAGGAT CTTAGACGGT TTTTGAAAGT | 12300 |
| CATrTTTCTT CATTAGACTT CCACAATCAG GGCAAGATGG AGCCTCATAA TCCAGCTTAG | 12360 |
| CGATAATTTT TTTGTGGGTA TCCATATTGA TGATATCTAG AATCTTGATG TTTGGGTCTT | 12420 |
| TAATATCGAG CAGTTTTGTG ATAAAATGTA ATTGTTCCAT ATGATTCTTT CTAATGAGTT | 12480 |
| GTTTTGTCGC TTTTCATTAT AGGTCATATG GGACTTTTTT TCTACACAAA AATAGGCTCC | 12540 |
| ATAATATCTA TAGTGGATT ACCCACTACA AATATTATAG AGCCCCAAAA GGAAGCCCTT | 12600 |
| TATGAATTGT AGGACTTCCT TTTCTTATCC AGAAATTGAT CTAGCTCTCT CTGATTTTGA | 12660 |
| AGAATAGTGA CTTTATGTGA ATATTCTTGG CAAAGTTTTT GGTAATTTTC TTTTGTAGTT | 12720 |
| TTGCGGACGC CCATCCCAA GAATCCATCT GATAAACTCC CACTCAAAGC GTTCAGGGCA | 12780 |
| ATCTACCGCC ATACTTTCTC TGACTTTTCC ACGGTATTTA AGATAACGCT TAAAGGCTCT | 12840 |
| AAAGAGACAG GTCAATGGCG AAAAATGAG AAAGATGATT TGGTCAGCTT CTTGCATTTCG | 12900 |
| TTCTTGGTAG TAGCACCAAG AATAATTACC ATCGATGACC CAAGCTTTAT GCTTGGTGAG | 12960 |
| AAAGTTTTTT ATCTCGGTTA ACATCCATTC GCAGTCACTG TCTTGCCAAC CAGGTTGAAA | 13020 |
| TTGGAGTGTG TCCATGTGCA GTTTTGAAT GGAGTAGTAG TTAGATAACT TTTCTGCTAT | 13080 |
| AGTTGACTTA CCAGAACCAG AATATCCGAT AATTGCGATT TTCATTTTCT ACCTTTTCCT | 13140 |
| ATTTGGAGAC AAAAAACAG CCTCTATGGA CTGTTTCTTA TTTAACAAGT TTAGCTGAAA | 13200 |
| GACGAGCTTT ATCGCGGCTT GCTTTGTTT TGTGAATCAA ACCTTTAGTT TCTGCTTTAT | 13260 |
| CGATAGCTGA GCTAGCAGCA CGGAAAAGTT CTTCAGATGG GTTTGCTTCG AAAGCTTTTA | 13320 |
| TAGCAGTACG CATAGCTGAT TTTTGAGCTG AGTTCCTTTC GATTCGTCTA ACGTTCAATT | 13380 |
| CAGCGCGTTT GATAGCTGAT TTAATGTTTG CCAATGGTCT TACCTCCATA TTTACTAACT | 13440 |

(2) INFORMATION FOR SEQ ID NO: 129:

890

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 8512 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 129:

| | |
|---|------|
| CCTTTTTTCA AAAACTAGAT ACTAGTCTAT CAAAAGTAGG AAAGGGTTTC AAGAAAATTG | 60 |
| ATTGGAAATT TTTTGAAAT CATAGAACTA TTAGCTAATC CCTAGTATTG AAAAGACTGG | 120 |
| ATAGCTTCTT TCAGGTCATC TTGTAACTA TTTCTCTGGT CAAGTTGGAC ATAGACTTCC | 180 |
| ACCAGACAGG ATCTAAAGTT GGAAAATTG TAAAAATCCT CCCTTTCTTC TATCGGAAAA | 240 |
| TCAACAGTTT TTATCCAAGA AGCTACTTGT TCTTGCTCCA ACTTCCCTTG TAAAATAGGT | 300 |
| TCATAGATCA CTCTTGCTAA ACGCCAATCC TCATCATCTG TAAAGCGAAT CGACATTCTT | 360 |
| TTAAATAGTT GGCCAAGTAT ATCAAATACT TCATGAACTC TGTMTTATAG AAAGTCTGGA | 420 |
| TGACAAACCA CCTCTGTCAG TAAATCGGCT CCATGTGCAA AAGCGTGAAC CCAACCATAC | 480 |
| TGACTTGAGA AACCCCTTGT ATCCTTTTCT TTTGAAAGAT AGTGCAAGCC TTGATTTAAA | 540 |
| AGGACATTAC GAATTTCTGG AGAAGGATT CCCAATGAT CAAACAACCA CTGGATTCTT | 600 |
| TCCTGGTTAT AATTTGGTTT TTCTTCTGCT ATTTTCTCTA GTAAATCTTG ATACATGGTC | 660 |
| AATACCTCTA CATTTCTAGC AACTGTTCAA AAAGGCAGTC TTAAATGACT CAATATTGAA | 720 |
| TTCTCAATTA AATACAATCT GATATAAAT GACGTAAATA ACTATCAATA CCAGTTCTAC | 780 |
| AGTAAGTTCA AATTTAACAT CACGACCTTC AACGACATT TTGAAAATAG CTACAATAA | 840 |
| GACAAATAGA ATGACGCTTA ACAAGCCCAT AAACATCATT CTAATAAATT TTTCTATTCC | 900 |
| CCTACTCTCC CAACTCAGCA CTATAGGAGA TAATCTGGTC AACTGTGTCA GACAAGAATT | 960 |
| GGATGGTATC ACGGAGTGGT TTGTCTGTTG AAATATCAGC ACCGATAATC ATGGCTGACT | 1020 |
| CAAGTGGTGT CTGCTACCA CCTGATTGGA GGAGATTGAG CCAGTCTTCA GCTCCAGTTT | 1080 |
| CAGAAATGTT TAGATGAAGG TAACCAGCAG TCGAGATAAC TAGTCCTGCT GAGTAAGTGT | 1140 |
| AACTATACAA GCCCATATAG TAGTGAGCTT GCGCATCCA AGTCAGAGTT GCATCATCGT | 1200 |
| CAATTTCAAT AGCATCTCCC CAGAAATCCG TCAAACTTC CTTCATAATG CTGTTGAGCT | 1260 |
| TGCTTGCTCC AAAGGTCTCC CCTTCTTCAA TCAATGTATA CACCTTACGC TGAAGGCGG | 1320 |
| CTTCCAAGAG GTGGGTGATA AAGTTATGGA AGTAGGTGTC TGTCAAGCGA TGAGCCAGAG | 1380 |
| CGAAGCGTTT TTGACGTGGG TCATTAGACT GGTTCPCAA GTAATCACTG AGTAGCAATT | 1440 |
| CATTGAAGGT TGACGGTGCT TCAACATAGT AGGTCGACAT ATGGGCATTG AAGTAACTTT | 1500 |

891

| | |
|---|------|
| GATGATTGTC TGAAAAGATG AATTGACCAG AATGCCCGAT TTCATGAATC AAGGTATAGA | 1560 |
| CATCGCTCAA ACGGCCTGTC CAGCTCATGA GTACATAAGG GTGTACGCGA TATGGGTCCG | 1620 |
| CCGCATAACC ACCGGAATCC TTGCCACTGT TAGCAGCAAA GTCCACCCAG CGCTCTTCTT | 1680 |
| GGTAACGAGC AACTTCCTGA CAATATTCTT GCCCAAAGG TTCTACCGAC TTCATGACCA | 1740 |
| AATCATAGGC ATCGTCAATA GTCAC TTCAG GATTTCAGGGC GCTGTCCAAG TCCAATTTCC | 1800 |
| AGTCTGCAAA GGTTCATCTTT TCAAGACCAT TTACCTTGGC AACATGCTTG AGGTATCTCT | 1860 |
| GAGCGACTGG TGCAAAGTCC TTCATGATGA GGTCAATCTG GCGGTCAAAC ATGACACGGT | 1920 |
| CCACTTCTTG TTCAGCTAGA AGATAGTCAA AGACAGAGTC GTATCCCTTC ATATCAGCCA | 1980 |
| AGAGTTTTTC AGACTTGACC TGAGCCAGAT AGGCTGCTGC AGCCGTATTT TGGTGCTTAC | 2040 |
| GAAGTCCCTC TGAGAAGGAA CGGAAGGATT TCTCAGCAAC CTCAGCATCC TCATGGTTTT | 2100 |
| GGTAGAAATT CTCATAGGTC ACAAAGCTGT TTTGTAGGT CTTGCCATGG GCTTCAAAGT | 2160 |
| CAGCCATTTT AAAATCCCCA GCTCGCATCT TAGTATAAAT GTCCTGCGGA CTGTAGAAAA | 2220 |
| CTTCACCGAG ATTTGTCAAG GCCTTCTCCA CATCTGCCCC TAAGTAGTGG GCTTTTTTGA | 2280 |
| TTTTAGCCTG ACGAATGGCA GCTGTTAAAT GTGGCAATTT ACCCAAACGG TCCAAGACTT | 2340 |
| CCTCATCTGC TGCCACCAAG GCATCGTCAA AGAAGGTCAA GGCTACGCTG GCATCTGTTT | 2400 |
| CAAATTCAT CCCAGCTTGG GCAATATGG CAAATTCGTC ATTGCTATAG TCCGTCGCTC | 2460 |
| GAGGCATAAA ACCATAGTTG CCAATATGGC TCATCTGAAT GTAGATCTGT TCCAATCCG | 2520 |
| CAAAGGCCTT CTCGAAATCC TCAAAAGTGT GAAGATTGCC CTTGTAATCA CGGCTAAACT | 2580 |
| GGTTGATGTC TTCGCGAGCT TTCTCGATTG CACGCAAGAA ATCCTCACGG TCTTGGTATA | 2640 |
| GGGCTGTTAA GTCCAGAGT TCCTTCTCTG GAAATTCGA ACGGTGTTTT TGTTCATTT | 2700 |
| TCTTCCTCTT ATTTCTCTAA TTCTACTAAA AACTAAGGG CTGATAAAGC GTAAAGCGGT | 2760 |
| GCTGTTTCTG CTCGCAAAAT ACGAGGACCT AGGCCTGCCA AAACGGCTCC TTTAGCTTCA | 2820 |
| AAACTTTTGA TTTCTGCAGG TGAGAGACCG CCTTCTGGAC CAAAGATAAA GAGCAGTTTG | 2880 |
| GCTCCTGTTT CAAGACCACT GACTGCTTGC AGAAGCGCAG CGGCTTCTCC TTCTTTAGCT | 2940 |
| GATTCTTCAT AGGCTACTAT GATAGAGTCA AACTGGTCCA GCTGAGCTAG AAAATCTGCT | 3000 |
| TTTTTCTCGA AAAGTTTAAT ACTTGGTACA ATATTACGCT TGCTTTGCTC GGCTGCTCCA | 3060 |
| AGGGCAATTT TTTCTAGTTT TTCAACTTTT TTACCCAATT TCTTGCCATC CCACTTGGCA | 3120 |
| ACTGACCAGT CTGCAGGAAA GGCCAGATT TGGCTAGCCC CCAGTTCGGT TACTTTTTGA | 3180 |
| GCGATGAACT CCAGCTTGTC TCCCTTGGGA AATCCAGATG CGATGGTCAC TTGGACTGGT | 3240 |

| | | | | | |
|-------------|-------------|------------|------------|-------------|-----------------|
| 892 | | | | | |
| AGTTCCACAT | TGTCATTTAA | TTCTTGGACC | AACTCAAAC | GACGATTTTC | CATATCCAGC 3300 |
| ACGCGCGCCA | AGCGCTTGAT | GCCATCATCA | AAGACTAAGG | TAACCTCATC | CTCTTCTTTC 3360 |
| AAGCGCATAA | CCTGAAACAT | ATGCTTACTG | GTTCCTTGT | CCTCGATAGT | GACAGGAGAG 3420 |
| ATAGCACTGC | CTTTTACAAA | ATACTGCTGC | ATGCTAGCCT | CCAATCACAC | CAGAGATATC 3480 |
| CTTGGTTTTTC | TTAAAGACAC | AGGTATTCCA | TTCCCTTGA | ACCATGTGAG | TTTCGAGGAA 3540 |
| AAATCCAGCT | GA CTCAGCCG | ACTGGCGCAC | CATGTCCAAC | TTGTCCCTGA | TAATGCCACT 3600 |
| CATGATCAGG | TAGCCTTCAT | CCTTTACCAA | GCGATAAGCA | TCGTCTATTA | GATGAATGAG 3660 |
| GATATCCGCC | AAGATATTAG | CCACAATCAC | ATCTGCCTCA | ATTTCCACAC | CCTTAAGCAA 3720 |
| ATCTCCAGCC | GCTACATGGA | TATTTTCCAT | GCCAGGGTTG | AGCTCAATAT | TTTCCTGAGC 3780 |
| CACACGAACC | GCCACATCAT | CCAGGTCATA | GGCGAAAATT | TCTTTAGCCC | CCAGAAGCGA 3840 |
| GCTGGCAATA | GAGAGAACCC | CTGAACCAGT | CCCCACATCT | AGCACCGTTT | CGCCACCACG 3900 |
| AAGAACCTGT | TCCAAGGCAA | AAAGGCTCAT | CTTGGTAGTT | GGGTGGGTTC | CAGTACCAAA 3960 |
| AGCCATGCCA | GGATCCAGCT | TGATAATCAT | TTCCCCGCA | GTCGCCTCAT | AGTCTGTCCA 4020 |
| AGAGGGAACG | ATGGTCAAAT | CATGAGTGAT | ACGAGCAGGT | TCATAGTATT | TCTTCCAGTT 4080 |
| GTCTGCCCAG | TCTTCCTCAG | CCAAGGCAGT | CGTACCTATT | TTTAACTCTC | CCAAATCCAT 4140 |
| AAAATCTGTC | AATTC TGCTA | GACGAGCCTG | CAAATCCGCC | TCAACCACTG | TCACATCCAC 4200 |
| CGTGT CAGGG | TAGTAGGCTG | TCACTACGAT | TTCTTCTTGC | TGCTCCACCT | CTGGGAAAAT 4260 |
| CTCTCCAAAG | CGGTCCACAT | TTCCACATA | GTCCATACTG | TCTTCGATTG | CGACTCCTTG 4320 |
| CGCTCCCAGC | TCAATCAAGA | GATTGAAAC | CAACTCCTCT | CCCTCACGCT | TCACTGTAAC 4380 |
| TTTAACTCT | TGCCATGTTT | CCATTATTAA | TACCAAGCCC | GTAAAACACA | AAACCAAAAT 4440 |
| AGGAAATTCT | CTGAAGACGC | TTGTGTCTAA | GAGAAGTTTA | TCTTTTGGC | ACAGTGTTTA 4500 |
| GGGCGGGTTC | AGTTTAGAAA | TGTAAGTAA | CCATCCTTTC | TAATCACTTA | CTTTTAAATA 4560 |
| ATCTTTTAAT | CTCTCTTGCA | ACTGAGGCAC | AACTTGACTG | GAAC TAAGAA | ATTCTTCAAC 4620 |
| ATTCATCAGC | TGATAGCCCT | GTCCTTCATC | TCCGAAGATG | ATATTGTCAA | ATTGTTCTTG 4680 |
| TCTTAGCTGA | CCAACCATAA | AGACCGATTT | CTTGCCTTTA | AAAATTACGC | TAGGATAAAT 4740 |
| CTTGCTCCAA | AGCAGACAGT | CTTCATCTAA | ATGAATTCCC | AGTTCCCTCAT | AAACTTCACG 4800 |
| CCGAGCGCAT | TCAAAAGGGC | TTTCGTCCCC | TTACGGCCA | CCACCTGGCA | GTTCCACAT 4860 |
| ATTGCCCCAG | GGAATACTTG | CCTTATCATC | GCGTAAGATA | GTCAAAAGCT | TATCCCCACA 4920 |
| AAACAAAGCA | ATCTTGCAAC | CTGTGAAATC | AGAAATTCT | AGTTCCATCT | TCAGTTCTCT 4980 |
| CTAACATTTT | CTTTTCCAGC | TCGGCTAACC | AGTTTTCATA | ATATCTTTTC | TCATCCCTCA 5040 |

893

| | | | | | | |
|-------------|-------------|------------|------------|------------|------------|------|
| ACATTTCGACT | ACTATCCATT | TTCTGTCTAG | CAATCTTGAG | AGCCTTACGA | GTTCGATCTA | 5100 |
| CATCTTTCTT | CACCTTTAAT | TGATACCAGG | CTTGTATCAC | TTGAAGATTG | GACAGTTTGA | 5160 |
| GAGACAGAAA | CGATTTGACC | TGTCGAATAC | TAGCATATTG | CTCCGCTTGC | TCAAAATCTC | 5220 |
| CTTCCAACAA | GGCGATATGA | AGCAGGGATA | GTTGGGCAAC | TGTCTGCATC | ATCGGAGTAG | 5280 |
| TTGTCCCTCTC | AAGTAATGCT | TGAAACTGCT | GTTTAGCTAC | TTCTTCCTTC | CCTTCCAAAA | 5340 |
| TGGAAACTTC | ACCTTGACATA | CCTAATACAC | CATCCGCAAA | ACTCCCTCGT | GCATCCTCAG | 5400 |
| GAAGTGTCTG | AACAAAGTCT | TTCAAATCAT | ATTCTTGAGG | AGCTAGCAAG | GTCTGGGCAG | 5460 |
| AATGTCTCAA | TACCAGGTAG | GCGTATTGCG | TATTTTCAGG | GTGTTGTAGT | AATTCCTCAA | 5520 |
| TTTTTGTCTC | ATCGGTGATG | TCGACTGGCA | AAATGTTATT | TAGGAAGAAA | GATAAATTAA | 5580 |
| GAAAAATCCA | AGTCCCTGCA | AAATACCAGC | TTCTTGTCAA | AAATCCAAAC | AATATCGCCA | 5640 |
| ATAATATCAA | GCCGAGATGA | ACCATCAAGC | CTCCTGAAAG | CATCAGGATG | ATTCTTTGAT | 5700 |
| CGCTTTCATC | CTCTTTTAAA | CCAATGTATT | GAGCACCAAC | ATTTTTCAGA | ATGGCTGTTT | 5760 |
| TACTAAGATG | AAACCTGCCT | GACTTTTGGG | TCAAAATAAA | ATGTCCTAAT | CCAAAAGCCA | 5820 |
| CCAGCCGATA | GCCTGTCAAG | TAGCCACAAA | AAGCATGACC | CAGCTCATGA | AGAATAAAGA | 5880 |
| TTAAATACAT | GCTTAGAAGA | GCGAAGGCAT | AACCAAAAGT | AAAGGCTAAA | ACTGCGGAAT | 5940 |
| ACCCCAACTC | TGCAATGCGG | ATTGTTCCAC | AAGCAAAAGC | TAGCATAATA | AAGACAACAG | 6000 |
| CTAGCACATA | AACCAATAAA | GTCCCAATTT | TCTTCATAAC | ACCTCCAACC | AACTCCTAGT | 6060 |
| ATCTTGGATA | AGGATAAAAT | TCTCCCTTTT | CCAAGCCAAT | TTTTCTCTCT | TCAAAGACTT | 6120 |
| CTTGGTTCCA | TTCCATGACA | AATTCCTCTG | CTTCTGGGTC | TTCCAAAAAG | TCCATGAGGA | 6180 |
| CATCTAGCCC | AACCTCAGCA | GTATCTTTAA | GGAAAAGCGC | AAAATAAGCT | AAAAATTCAC | 6240 |
| GGGAAAATCC | TTTTTTAGGC | AGGTAAGGAA | TAACAGTCAA | ATAGTCTTCC | TCATTGACTG | 6300 |
| TTGACTTGGC | AGGATTGTAG | AAAAGGACCG | CTTCCTCAAA | AAGAATGTCA | TCTGATGAAA | 6360 |
| CCTCTCCGTC | TTCATCCACC | ATCTCCACAC | CGCAGCATTT | TGCGCTTCCA | ATAGAAAAC | 6420 |
| CACTTCTACC | GCATGGTTGC | GTTTGTCCCA | GCTAATCTCA | AAGTCAAAGG | GAAAGTTCTT | 6480 |
| GTCCAACCTC | TCCTCTAAAA | TATCTAAAAA | TCCGTATGTT | GCCATTTTGT | CCTCTTTCTA | 6540 |
| TGCGACTCTT | TAATCGCCCC | GATTGCTCGG | AAATATGCTA | AAATAGATAC | TACCATCTTA | 6600 |
| CCACAAAATT | ATTTTATGTC | CTAATTATAC | CATATTACCT | CATTTAAACC | CTTGGTATCA | 6660 |
| GTGATTTTCT | TAAAAGTCTG | ATTTCTTCAT | TTCTCATAAA | AATCAATATA | AAAAGCCCTC | 6720 |
| GAAAGGGCTA | ATAAATCTAT | AAAATCAATA | GGCGAGTAAC | TAGCACAAAG | GGACGTGCTT | 6780 |

894

| | |
|--|------|
| TTTTATTGAC TATTACCACG ATACCACGCT TAATCTTAGG CTTGAACTTT CTTATCTGCA | 6840 |
| ATAGCGTCTG TCAAAGTCTG AGAAAAGTTA AGCCCCATTT CTCGTCCCAA CTTATCTGCC | 6900 |
| CATTTTGGTA TGGTCAAAGT CTTTTTAATG GGTTCCTGAC TTCCTAGGTA TTCTGATACA | 6960 |
| TCAACAGATA CCATAGAAAT AAAAGATTTA TCAAGGTCAT AGGTTGACAC GAAATCTTCA | 7020 |
| TCATCTTTAA AAGGATCATT ATCAATTAAA GACAAGCTAT TGATATCTGA TGGCTGAGGT | 7080 |
| AACTCTCCAT CACTCTCTAT CAAATCTGCA ACAGTTATCC CTAGCCACTC CGACCCATA | 7140 |
| GCCAAAGCCT CAGAAATCCC CTCTCCTTGT GTAGCTGAGT ATTCAAAATC TGGGAAATGG | 7200 |
| ACAAAATAAG TCGCTTCTGT TCCGTCTGTG TCGTCATAAT AAAATAAAGC TGGATACGTA | 7260 |
| ACTAACATTT CACTACCTCC ATATCAAAAA GCAGGGACTG AATTTTACAA CCCAGCTTGC | 7320 |
| TTTCTTATCC CTCTTTCAGT GTACTTATTC AGCTCACCAT GAAGGATTGT GATAGGTCTT | 7380 |
| TCCCCTTGCT TTTCCATTTT AATATGGGAG CCTTTACCGC CTCTAGTCTT TATCCAACCA | 7440 |
| TGGGCCGTAA GGAGTTTAAC CATCTCTTTT TGTGTCATAG GCATAGCGCT TTTACCTCCT | 7500 |
| GACAACACCA TTATAACACG TGTTACACGT ATTGTAAAGG AGTGATACTT ATTATTCTAT | 7560 |
| TATACATAAA AGCCCCTAGA TGTGGTTCTA AGGGAAGCCA ATTTATTTCAT ACCTATTTTT | 7620 |
| CTAATGAGTA GTAAAACTG CTTCTTTATC GAGCAATTCA TCATCTGTAT AGTCAATTGT | 7680 |
| AAAAGTATCT CGATCTAAGA CAGATTGAGG CGGAGTTGAA TGAATCATAG GAACACTGCG | 7740 |
| TACTCTATAT TTTTATCTC CAATTTTAC AAAGTATAC TCTTCGAAAA TCAAATTCAA | 7800 |
| ACCACGTCAA CGTCGCCTTA CCGTACTCAA GTACAGCCTG CGGCTAGTTT CTTAGTTTGC | 7860 |
| TCTTTGATTT TCATTGAGTA TGATTAACTC TCAAGTCTTC GAAATCAGGA TTTTCAACAG | 7920 |
| TTATTACAAG GAGGCGATTT ACTACTTCAA AAACATCAAT TATTCTATTT TTCATATTTT | 7980 |
| TTCAACCCAT TATTAGAATG AACTTCTTGG TAAGCAAAAT CAAGTTTAGA TTTAATGTTT | 8040 |
| TCGTACAAAT CTAAAATCTC TTTTGGAGTA TCTTCCCGGA AGAAAAGTTT TCTTTTCCCT | 8100 |
| GAAATAACTT GATCACTAAG AATCCAATGA CGAATTGTGTT TTGTAAAAAT CAAAATTCC | 8160 |
| TGACTTGGTA GTTCCATCAT TTCCATTGCT TATCACCTCT CTTTTCATTA TAGTTCATAC | 8220 |
| AATGACATTC AGCAATATTA TTTCTCAAGT CAGCACTTCC ACTTCTTTAG GCTCAACTAT | 8280 |
| CCTATTTTGA GCTTTAAGGA AAATCAAATC TCTCATGCTG ATACCTCTCC TCATTAAATT | 8340 |
| AAATAGTAAA AAAGATTCTA TCTCACTCCC TGATTATTAC AAAACCATTG AAATATCACA | 8400 |
| ACTAATAGGC TAGAATGGAC ATAGTAAGAT ATAGTAGATG AGTCATTCTA CTCAAATCCA | 8460 |
| CGTTAGAAAG GACTGCTATG CCAGACAATC TCGCCGTTCTG CATGCGCCCG GG | 8512 |

(2) INFORMATION FOR SEQ ID NO: 130:

895

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2869 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 130:

| | |
|---|------|
| CTCGTTTCAA GGTGAGTCT CTTGCAAATC TTGTTGCGCT TCTTCCTTTT GCCAAGGCAT | 60 |
| CTCTCCCATG GTTGGTGCCa GCCATTGTTG GAATCTTGCT CTCATTGGTT CTACCAAACA | 120 |
| AGCAAGAAAG CGATGTTTTT GAAATGGAAT AATCACTTAA ATCACTTTTG TAGCCAAGTC | 180 |
| TACAGGAGTG ATTKTCTTTT TTTATCCGAT GATAAATGTG TTATAATAGG TAGCGAAAGA | 240 |
| GGTGAAGAAA TGAATCAAAC AGTAGAATAT ATCAAAGAAC TGACAGCCAT TGCGLCGCCA | 300 |
| ACAGGCTTTA CTCGTGAGAT TGCGGACTAT TTAGTCAAGA CTCTAGAAGG TTTTGGTTAC | 360 |
| CAGCCGGTTC GCACATCCAA GGGCGGTGTC AATGTAAC TAAGGTCA AAATGATGAG | 420 |
| CAACATCGCT ATGTGACTGC CCATGTAGAT ACGCTTGGTG CTATTGTCCG TGCTGTCAAA | 480 |
| CCAGACGGCC GTCTCAAAAT GGACCGTATC GGTGGCTTTC CTGGAACAT GATTGAAGGA | 540 |
| GAAAACTGTA CCATTTCATGT GGCTAGCACA GGTGAAAAG TATCAGGAAC CATCTCATC | 600 |
| CACCAAACCT CTTGCCATGT CTATAAGGAT GCAGGAACG CAGAACGCAC GCAAGACAAT | 660 |
| ATGGAAGTGC GTTTGGACGC CAAAGTAACT AGTGAAAAAG AAACCTCGTG TCTTGGCATT | 720 |
| GAGGTCGGTG ATTTTATCAG TTTTGACCCA CGAACTGTG TGACAGAGAC AGGTTTATC | 780 |
| AAGTCTCGCC ATTTGGATGA CAAGGTCAGT GCGGCGATT TGCTCAATCT CCTTCGCATT | 840 |
| TATAAGGAAG AGAAGATTGA ATTGCCCGTA ACAACTCATT TTGCTTTTTC AGTCTTTGAA | 900 |
| GAAAGTGGAC ACGGTGCAAA CTCTAACATT CCTGCTCAGG TAGTAGAATA TCTGGCTGTG | 960 |
| GATATGGGAG CCATGGGAGA TGACCAGCAA ACAGACGAAT ATACAGTGTC TATCTGTGTC | 1020 |
| AAGGATGCTT CTGGACCTTA TCACTATGAC TTCCGTCAAC ATTTGGTGGC TTTGGCGAAA | 1080 |
| GAGCAAGATA TTCCATTTAA GCTGGATATC TATCCATTTT ATGGTTCGGA CGCTTCAGCG | 1140 |
| GCTATGTCTG CAGGGGCAGA AGTCAAACAC GCCCTTCTCG GTGCTGGTAT AGAGTCTAGC | 1200 |
| CATTCCTATG AGCGTACCCA TATTGACTCG GTGATCGCAA CAGAACGAAT GGTGATGCT | 1260 |
| TATCTTAAGA GCACGTTGGT GGACTAATAT GTGCCTTATT TGTCAGAGAA TTGACCTCAT | 1320 |
| CAAGAAGGAA GAAAATCCTT ACTTTGTCAA AGAGTTGGAA ACAGGCTATC TTGTGGTTGG | 1380 |
| AGACCACCAG TATTTTGAAG GCTATAGTCT CTTTCTAGCC AAGGAGCATG TCAGCGAATT | 1440 |

896

| | |
|--|------|
| GCACCATTTG AAAAAGGAGA CAAGACTCCG TTTTCTAGAA GAAATGAGTT TAGTCCAAGA | 1500 |
| GGCAGTTGCC AAGGCCTTTG CTGCTGAGAA AATGAATATC GAACTGCTAG GAAATGGCGA | 1560 |
| TGCTCATCTT CATTGGCATC TGTTCACACG ACGGACAGGT GATATGAATG GTCATGGTCT | 1620 |
| CAAGGGTCGT GGACCAGTCT GGTGGGTTC CTTTGAAGAA ATGACAGCAG AAACCTGCCA | 1680 |
| AGCAAAACCG GATGAGATTA AAAGATTAGT CAAACGTTTA TCGTCAGAAG TAGATAAACT | 1740 |
| ATTAGAAATA AAGGAGTAGA AATGAAGAAA AGATACCTAG TCTTGACAGC TTTGCTAGCC | 1800 |
| TTGAGTCTAG CAGCTTGTTT ACAAGAAAAA ACAAAAAATG AAGATGGAGA AACTAAGACA | 1860 |
| GAACAGACAG CCAAAGCTGA TGGAACAGTC GGTAGTAAGT CTCAAGGAGC TGCCCAAG | 1920 |
| AAAGCAGAAG TGSTCAATAA AGGTGATTAC TACAGCATTC AAGGGAAATA CGATGAAATC | 1980 |
| ATCGTAGCCA ACAAACACTA TCCATTGTCT AAAGACTATA ATCCAGGGGA AAATCCAACA | 2040 |
| GCCAAGGCAG AGTTGGTCAA ACTCATCAA GCGATGCAAG AGGCAGGTTT CCCTATTAGT | 2100 |
| GATCATTACA GTGGTTTTAG AAGTTATGAA ACTCAGACCA AGCTCTATCA AGATTATGTC | 2160 |
| AACCAAGATG GAAAGGCAGC AGCTGACCGT TACTCTGCCC GTCCTGGCTA TAGCGAACAC | 2220 |
| CAGACAGGCT TGGCCTTTGA TGTGATTGGG ACTGATGGTG ATTGTTGAC AGAAGAAAAA | 2280 |
| GCAGCCCAAT GGCTCTTGA TCATGCAGCT GATTATGGCT TTGTGTCCG TTATCTCAA | 2340 |
| GGCAAGGAAA AGGAAACAGG CTATATGGCT GAAGAATGGC ACCTGCGTTA TGTAGGAAAA | 2400 |
| GAAGCTAAAG AAATGCTGC AAGTGGTCTC AGTTTGAAG AATACTATGG CTTTGAAGGC | 2460 |
| GGAGACTACG TCGATTAAATA CTCTTCGAAA ATCTCTTCAA ACCACGTCAG CGTCGCCTTA | 2520 |
| CCTACTGACT GCGTCGGTTC TATTCACAAC CTCAAACAG TGTTTTGAGT cGATTCTGCA | 2580 |
| GTTTTATCTG CAACCTCAAA GCTGTACTTT GAGCAsTGCG GCTAGCTTCC TAGTTTGCTC | 2640 |
| TTTGATTTTC ATTGAGTACA AAAAGTAAAC TTTTCTCTTG CAATTCAGTA TAAATAGTGT | 2700 |
| ATAATGGATG GGTATGTGAA AACATACTT GTGGGAGGTA AAAATCTCTA ATTACCGCCA | 2760 |
| AAACCACAAA GGAGGATTTA AAAATGGCTA AAAAGTCGA AAACTTGTA AAATTGCAAA | 2820 |
| TCCCTGCTGG TAAAGCTACA CCAGCTCCAC CGGTTGGACC TGCTCTTGG | 2869 |

(2) INFORMATION FOR SEQ ID NO: 131:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 6186 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 131:

897

CTGAATCCCT TATAGGAGTC CAGTAACTTT TTAGCCTCTA CTTTGCCTTC ATAGGCAGCT 60
 TCAACATCAT TAAAAAAGA ArGCACTGAA GCAAGTTCTT CAGTGCTCCA CGACAAATCT 120
 AGTGGGTAAC TATACTGTTT GTTCATTAAC TAATACCAGC TCTCATTTCTT GCTTCTTTTA 180
 GTTCTTGCTT ACGATAACTA CGAGGGAGAA AAGCACGAAT CTCATCTTCA TTAAAACCGA 240
 TTTGCATACG CTTGGCATCA ATAATAATG GACGACGCAA AAGACTAGGA TACTGCTCAA 300
 TCAAATGAAG CAATTCCGAT ACCGAAATAC TCTCTACATC AATATTCAAT TTTTGAAAAA 360
 TTTTGAACG AGTTGAAATG ATGTCATCAG TACCATTTC GGTCAAGGAA AGGATGTGTT 420
 GCAATTCTTT TCTTGTTAAA GGACTGGTCA TAATATTGTG TTCCACAAAG GGAACCTATG 480
 TTTTCTAAC CAGGCCTTAG CCTTACGACA TGATGTACAG CTCGGTGATA GAAATAGTGT 540
 AATCATGCTT TTCTCTTCTT ATCTATACTT TGCTACTTCT ATTATACAAA AAAATAAAGC 600
 GCTTGACTAG GGATTTTTAG AAAAAAGCC TATTTTTTCA AGAAAAATAG GCTTTTTGCG 660
 AACGATTGAC ACAATTGGAT TTGGTTAATT CACTCTTAAC GATGGTTTTA AACGATATAT 720
 ATTTTTATAT ATGTAAATTA AAAACATCTT TCCTTTCACT TCCTACGACT TTTCAGATAC 780
 AGATAGCCAA AGAAGTTTTC ATAGAGGGCA AAAAAGAGGA GGAAGGCATG AAGAAAGAAG 840
 GTCTCTGGCA AAATCATAAT AACAGGATCC TTGGCTGGAT CAAAAGCCA GGTATCATCT 900
 CCCACAAAGA GAATTTGATG GAAAAGAGTA AAGAATTGGT CAAAACCAAT CAAAACCTCC 960
 CCAAGTCCAA TCATCACAGG TAAGACTACT AGAGCCAGGA GACTTTTTTCG ATAAAGAGAC 1020
 AAAAAGTCCT TTTTCACAAT CCTATTGACA AAGACATAGA AACTTGGCAG TGTCACTAGA 1080
 GCTACTAGCT GAACCAAATG AAAGAGATTC TTGACCACTG CGAAATGGTG CAGACCAGCT 1140
 GCTGACGAAC GAAAATCAGG CATCTGTAAG ACCTGACTAA AAGGATTGGT CAGATAATTC 1200
 ATCAAGATAT GAAAATTGTA TTGAATGGTT TCTGGTTTTA GATAGACTCG ATTCTGTTAAG 1260
 TTTAGCCACT GAATCTCCAT AGGATAGAAA ATCCAAGCCA GATAAATGGT CAGAAGGATG 1320
 GAGAGGGAGA GGAGAAAGAG CATAGAGCCC CAAAAGATCA ATTTAGTTTTT CATCAAAATC 1380
 CCACTCCGCA AGGCTAGAAA CCACATGTGT CGGTGCGATT GGCAGGCCAG CTACTTCTTC 1440
 TGCCTTAGTA AAACCTGTCG TCACCAAGAG CGTTGGAATG CCATTGTCAA TCCCAGCCCC 1500
 AATATCAGTC AAATAATTGT CCCCAACCAT GATTAACCTC TCACGTTCCA AACCTAAGTG 1560
 CTCAACCGCC TTGTCCATAA TGATGGCATT TGGTTTTCCG ATATAAACCG GCTTCACTCG 1620
 TGTCGCTACT TCAAGCAGCG TAATCAGTGA GCCAGCACCT GGCAAAAGAC CGCGTTCCGT 1680
 CGGGATGTTG AGGTCAGGAT TGGTCCGAT AAAATGGGCA CCCTTTTGAA TAGCAAGAGT 1740

898

| | |
|--|------|
| TGCTGTGGCA AATTTTTCAT AGTCGACTTG CCAATCCAGA CCAACTACCA CGTAGGCAGG | 1800 |
| TTTTTCCTTG TCTTCCACAT AACCAGCCGC CTTGATGGCT TCCTTGAGTC CTGCTTCTCC | 1860 |
| GACGACATAG ACGGTCTTTT CAAGCCCCAA ATCATTCTATA TAGTCGATGG TTGCCAAAGT | 1920 |
| CGCTGTGTAG ACACTCGATA GGGGCGTATC GATATTAAAA TTCTGAGCCA ACATCTCCTT | 1980 |
| AACACTCTCT GGAGTGCGGG TTGTATTGTT GGTACAAAG AGATAGGGAA TGTCCCGCTT | 2040 |
| TTGCAATTCA TGAACAAAAG TCTCTCCAGC AGGGATTCCG TCTTTCCCCT TATAAATGGT | 2100 |
| TCCGTCTAAA TCAATTAAAT AGCCTTTATA TTTTCATCTAT TTCTCCCTAA GCCTTTTATA | 2160 |
| TTTCTTGCCA AGTAATGATT GCTTGGGCAT TGATAACCCC ATCACTTGTA ATTTTCATGCT | 2220 |
| TGCTTTCCAG TCCAGTCCGT TCAACAGCCG ATGTAATCAC CCCACCTGGT CGAAGTTCCCT | 2280 |
| TGACATACTT GAGGTGATT TTCTTGGGAA TATAGTGGGT CAAAAATCC GCTCCCATGA | 2340 |
| CCTCAAAAAT CCAGTCCAAG TATTTACTGT TATTGACATG ACCATTCTATA TCCAAGTCGT | 2400 |
| AAAAACGAAC ATGGTAATCC TTGCTGATCG GTTCTTCCAA GGACTCATAC TTCGGTCCAC | 2460 |
| GGATAAGTTT TTATCAAAA TCAGACTGGT AAGGAGCCAC AATCTCAGGT TCAACAACAT | 2520 |
| GGACTTTTCG ACTGTCGCGG TCCATGAGAA CAAAGGTCGC CATCATGTGG ATGAGCTCCT | 2580 |
| GCTCCGCTTC ATTATAAATA GTAAAGCGAC GGTCAGAAAA AAGTCGATTG TAGCTCAAGG | 2640 |
| CTTCCGTTTC GATGGTAATT TCTCCGCAA AACGAGGCAA ACGAACCACC TCAATATCAT | 2700 |
| ATTCTACGAT AATCCAGACC AGATTATATT CTTCCTCAAT GGCCTTATCA CTAACCTCCA | 2760 |
| GTTCAATCGA CTGCATCCCT GAACTTGCA GTGACAGCAA AATCACATCT GGAAGTTTGA | 2820 |
| TATGACCGTT CATATCAGCC ATATCAAAAG GAATTTTCAT TTTCATTGA TAAGTTAAGC | 2880 |
| CCATGATCCT ACTCCAAAAT AAATCGTTCT GCTACAGTAT CTCCCAAAA GAGACCTCTC | 2940 |
| TTTGTCATGC GAACGTGGTC ACCCTCAATC TGCATGAGGC CTTGTTGAAC CAAATCTCTG | 3000 |
| ACAATTTCTC CATAAAGTCC AGCAAAAGAC TGTCCAAAT TTTCTCTCAA TCGCGCCATG | 3060 |
| GAAACCCCGG ATTTCTTGCG GAGTCCCAAG AACATTTCTT CTTCATTG CTCTTTTGA | 3120 |
| CTCAGGTGAT CTTCTGTAAT ACAAGCATTG CTTCTCTCAA CCGCACTGAG ATAATGACGA | 3180 |
| ATGGGACCAT GATTTTATA GCGTACTCCA TTGACATAAC CAGATGCCCC TGCACCAATA | 3240 |
| CCATAGTATT CAGCATGTGC CCAGTACATG AGATTATGAC GACTTTTCAA ACCGGGTTTG | 3300 |
| GAGAAATTAG AAATCTCTATA ATGCTCAAAA CCCGCTCGCT CCAGCTCTGC AATGATGTAC | 3360 |
| TCAAACATCT CCGCTTCTAG TTCTCCTTA GGCAGAGGCA ATTTCCCACG TCGCATCCGG | 3420 |
| TTCATAAAGA CCGTATGGTT TTCTAAAATC AAACATATACA AACTCATGTG GGGAATATCC | 3480 |
| AATCCAATGG CTTTAGCCAC ATTTTCCTTT ACTTGCTCCA TGGTCTGACC AGGCAGAGCA | 3540 |

899

| | |
|---|------|
| TAAATCAAAT CAATGGAGAT ATTGTCAAAA CCAGCCAGTT TCAGGCGATC GATATTTTCA | 3600 |
| TAAATATCCT TCTCCAAATG ACTGCGCCCA ATCTTTTTC AATCTTTATC ATCAAAGGTC | 3660 |
| TGGACACCTA GCGAAACACG ATTGACAGCC GAATTTTTC AATCAGCTAT CTTATCCGCA | 3720 |
| TCCAAATCGC CTGGATTGGC TTCAATGGTC AACTCTTCCA AGACAGACAA ATCCAAGTTT | 3780 |
| TTAGTCAAGC CATTCAGTAA CACCTCCAGT TGCGGAGCCG ACAGGGCTGT CGGTGTTCCA | 3840 |
| CCACCGATAT AAAGGGTTGA CAACTTTTCA ATATCATAAG AACGAACTC TTCCAGCAGA | 3900 |
| TGCTCTAAAT AGCTGTCGAC TGGCTGATTT TTGATGAAGA CCTTTGAAAA ATCACAATAA | 3960 |
| TAACAAATCT GGGTACAAAA TGGGATGTGC ACATAGGCTG ACGTTGGTTT TTTCTGCATA | 4020 |
| GTAATTATTA TACCACAAAG ACTAGATTCC AGATAAAAA CACCATCCCC AGATACATAG | 4080 |
| TCCGTCGGA GATGGTGATG GTTATTCTT CTGTTATATC AATCACAATC TCTTCTGAGT | 4140 |
| CATCAAGAGC TTCGGCTTTT TCTTGCCATT GCTCCTTGAG ATTATTTAAT TGATTTTTTG | 4200 |
| ATGCTTCTGT CGCTTGAAAA GCATAGGATT TAGTTTGAGC AAGTATACTG TCCACAGTGA | 4260 |
| TTTCACCTGA CTCAACCTGT TCTTTTGTTT TCAGAACAAA ATCTGTAGCC TGCTCCTTAA | 4320 |
| CTTCTGTCAG TTTTTCACAG ACTTGCTCCT TGGCATACTC CGGATCTTCT CTCAAATCAT | 4380 |
| CTAGAAAAATC TTGAGCCTGA CTGCAAACTT GTTGGCCCTT ATCACTTGTT AAAACAAGG | 4440 |
| CAAGAGCTGC ACCTGAAACG GTTCCTAAAA GGATTGAGGA TAATTTACCC ATAAGGATTC | 4500 |
| TCCTTTTTTA TTTTGTGAAA AATTTACTTG CAAGACGAAG AGCTGACAGA CTTGCACCAG | 4560 |
| TCTTGAGTGT TTTTGAACCA GCTGATGAAG CTTTCTTGCT CAAGACACGC GCATGGTCAT | 4620 |
| TGAGGTCTGA AACAGATAGA GATAAATCTG CAACAGCACT GAAGAGTGA TCAATCGTAG | 4680 |
| CCACCTTGAC ATTGATATCA TCTGCCAAGA CATTGACCTT AGCCAACAAC TCATTGGTGT | 4740 |
| GATGCAAGGT CACATCCACA TCTGAAGTCA AGGTTTAAAT CGTCTTTTCT GTTTCATCGA | 4800 |
| TGACACGACC AAGCTTTTGT ACAGTAATGA TCAGATAGAC CAAAAAGACA ATCAAAGCTA | 4860 |
| GGGCAACAAG AATATATGCA ACTTCTAACA TTTAGTTTTC CTCCTCTGTA ATATAGTAAG | 4920 |
| GGGCCTTCTT TCGATTTTGA TAAATAACGA TCATTATACC GAGACCGATA AGGACAACTG | 4980 |
| ACAGCCATTG GGACACTCGA AAGCCGAAGA ACATGAGACT ATCTGTTTCG ATACCTTCGA | 5040 |
| TAACCATACG ACCGAAACCA TACCAATCA AGTAAAAGGC CGTGATATGA CCTCGTCTGA | 5100 |
| GACTCTTCCA TTTCCGTCTA AAAATCAGAA TCAAGGCAAA GCCAAGCAGA TTCCATAGAG | 5160 |
| ACTCATAAAG GAAAGTCGGT TGACGGTAGC TCCCCTCAAT ATACATCTGG TCACGGATAA | 5220 |
| AGCCAGGTAG ATAATCCAGA TTATCCACTG TTGCACCATA AGCTTCTTGG TTAAAGAAAT | 5280 |

900

| | |
|---|------|
| TACCCCAACG CCCCAACTT TGAGCAATCA TAACGCTAGG CGCCGCAATA TCTAGAAAAT | 5340 |
| CCCAAGTATT GATGAGTTTA CGGTCAGCAA AGATATAGAG CACAAGAGCC CCAGTTATCA | 5400 |
| AACCACCGTA AATGGCCAAA CCACCATTCC AAATGGCAAA AATCTCTCCT AAATTCTGAC | 5460 |
| TATAGTAATC AAATCGGAAA ATAACATAGT AGAGACGAGC TCCTAAAATA GCCAAGGGAA | 5520 |
| AGGCTACTAA GATAAAATCT AAAATATCGT CTGGTATGAT CTTCTTTCTA GGTGCTTCTT | 5580 |
| TCATGGTCAA ATAAACCGCA AGAATCAAGC CTGTCACAAT ACATAAGGCA TACCAACGAA | 5640 |
| TGGCTAGGGG TCCTAGTTGA ATAGCAATTG GATCAAGCAT TTTGCACCTC ATTTGAGCGG | 5700 |
| ATTAGACTTG TCAGTCGTTT GTCGAACAAA CGGGTCGCAT CAAAGCCCAT TTCCTTGCCA | 5760 |
| CGATAATTCA TGGCAGCTGC CTCAATCACA ACAGAGATAT TACGACCTGT TTTAACTGGA | 5820 |
| ATACGAATAC GAGGAATGtA CGCCAGAAAC TTCAAGTTCC TCTGCATTAT TTCCAAGACG | 5880 |
| ATCAAAGGTC TTATGCGTAT CGTAATTTTC CAAATAGACA GCAAGCTGAA CCTGTGAAGA | 5940 |
| ATCCTTGACA GCACTCGCAC CGTAGAGACT CATAACATCG ATAATACCAA CCCCACGAAT | 6000 |
| TTCAATCAAG TGTTCAAAA TTTCAGCTGG TTCACCCAG AGAGTAATCT CATCCTTGGC | 6060 |
| AAAGATATCG ACACGGTCAT CGGCTACCAA ACGGTGACCA CGTTTGACAA GCTCAAGACC | 6120 |
| TGTCTCGCTC TTACCAATTC CACTATCTCC CTGAATCAAG ACGCCCATCC CATAAATATC | 6180 |
| CATCAA | 6186 |

(2) INFORMATION FOR SEQ ID NO: 132:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 9541 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 132:

| | |
|--|-----|
| GAAATCACA ACCCTTTTGG CAAAATTTT GAGATTATTT TCACAAACTT GATTTTCAA | 60 |
| AGTATACTCA ATAAAAATTA AAAAAATCCA CTACGTCAAG GCGAGGCTAA TGTGGTTTGA | 120 |
| AGAAATTTTC GAAGAGCGTG AATGAGTATC ATCTATAGTA AAATAAAAAA ACTGAACAAT | 180 |
| TTGGTTGGGG ACAGCCAAAC CAATTTCTCA CAATGTTTCA GAAACAAGGG TGTGCTATTC | 240 |
| CAATTTGAGC CTACTATAAC TGTCATAGAT TGCTGAAACA AAGTCTAGGT AAAAGTCTTC | 300 |
| ATAATAAAAA GACCTCCTAT CAAGTGTTCA AAAACTTTGA TAGGAGGTCT TGTTTTGTGA | 360 |
| AAATATTTAT CAAATTTTCT ATACAAGTGA GCTGTTAGCC AGGTTCTTTC TATTTCTTCA | 420 |
| ATTTCAATGA ATGGATTTTT TACTAATACT CATAACTGGG AATTTGCTCTG TGTAAAAATA | 480 |

901

| | |
|--|------|
| GCGAGATAGA TGGTATTTAT AAAACACTCA AGACAGCTAG ACTAATATCA TTTAAAACAT | 540 |
| TATCTTCTTT TGAGCGACTG TTGGTTACCA ACATAGCTAA ATTTCTTGCA TTTTCAAATT | 600 |
| GATAGGGTTC TGATTTAGCA TTCACAACCA CCAAGAGGTG TTCTTTGCCG TGAAC TTCAT | 660 |
| AGATAAGGTA GCCGCTATGT TCAATCGCAG AATGCACAAA GACATGATGG TAAATTTTCAT | 720 |
| CATAGCTAGA GTAAGAAAAG GCACCAAGTTT TTGTCTTCAA TCGGATGACT TGACGGATAA | 780 |
| ACTCAATACT GTCTTGACGC TCATTAATCA AGTTCCAGTT CACTTGGTTC ACAC TGTCAG | 840 |
| GAGCATTATA GCTATTCATC GCACGCTCTC TATCATCATG GGTCAACTCA CCATTTTCAC | 900 |
| CAGTCGCAAC CAGTTTGGTA CGACCAAATT CTTGACCGAT TTCCATAAAG GCCATCCCCCT | 960 |
| GCATGAGCAG ATTCATGGCT GTGGCAGTTT CGACCTTGCG CATGATTTGC TCTGAACTTT | 1020 |
| GGTCTGGATG AAGGGTTGCC AATAAATCGT GAAGATTGTA ATGTGCATGG GCTTCTACAT | 1080 |
| AGTTAAGCAC CTGATTTGGA TGTGTATAGC TTCCTAATTC ACGACTTCCT AGGATTGCTT | 1140 |
| TAGCTAGAAT TGGCTCTGTC GCAGCACCAC TGACAAAACC TGACTTGATA GCACCATAAA | 1200 |
| CTTCTCCCCC TTTGACAGCA TCGCGCTGAT TGTCAATAAA GAAACCAATA TTTGGCATCT | 1260 |
| GGTAGGCATT GTCCTTCTTG GCCTTATCAT AAGGGGCAAG ACCTGTTCCC ATATCCCATC | 1320 |
| CTTCTCCATA GAGGATAATG TTGGAGTCGA TTTTCATCAA GCTTTGACGA ATCATCTGCA | 1380 |
| TGGTCTTGAC ATCATGAATC CCCATCAAGT CAAAACGGAA GCCGTCAATA TTATATTCCT | 1440 |
| GCACCCAGTA TAGAAGAGAA TCAATCATAT ACTTGCGAAA CATTTCTGTG TCACTGGCTG | 1500 |
| TTTCATTTCC AACACCCGTT CCATTCTGGA AGGTACCATC TGGATTCATA CGATAATAGT | 1560 |
| AATCAGGGAC TGTGTTTGG AATGGTGCAT CAACAAC TGA GAAGGTATGG TTATAGACTA | 1620 |
| CATCCATAAT GACTCCAATA CCCGCATCGT GATAAGCTTG AACCATCACC TTCAAATCAC | 1680 |
| GAATGACCTG AGCTGGATCA TCTGGATTAG TTGAAAACT AGTTTCTGGC GCGTTATAGT | 1740 |
| TTTGTGGATC ATAACCCAG TTGTAGGTTA CATTTCCATC CTCATCGTAT TCTTTATGAC | 1800 |
| GGTCTGCAAT TGGTTGCAAT TGAACATAAT TGAGCCAG CTTCTTGATG TAATCAAAAAG | 1860 |
| CAGTTGACTG GCCGTATTGG TTAACGTTC CAGCCTGAGC AGCACCCAAG AAAGTTCTCTC | 1920 |
| GAAGATGTTT ATCTACACCC GATGTAGGTG ATTTAGTCAA ATCACGAATG TGCATTTTCAC | 1980 |
| AGATAACTGC CTTACATGGA TTTTCCAAGC GCCAAGTAGC CTCCGAACCG TGCTTAACCT | 2040 |
| CGAAGTTTTC AACTTGCTTT TCTACATGGC TCAGAATAGC TGAACGTTTG CCATCAGGGC | 2100 |
| TGGTCGCGAT TGTATAAGGA TCACGTGTCA GTGTTTGGTG ATGAGGGAAT TGGACTTGAT | 2160 |
| ACTGATAAGT CTTACCTACC AAATCTTCTT CAACATCCAA ACTCCAGACA CCGATTGTAT | 2220 |

902

| | |
|---|------|
| TGTCCTTATG ATTATAAGAG TAGCTATTGC CTCTTTTCAT CTCAAAAGTC TTCCAAACGG | 2280 |
| GTGCATCATT AGCAGCTGAT TCATAACGA CAACTTGAC TTCTGTGCGT GTAGGTGACC | 2340 |
| AGAGAGAAAA ATGAGCCTGA TTGTCTCTA CACGGCAACC CAATTCTCCT TGGTAACCCC | 2400 |
| AATGATGATC AAAACTAGCA CTGTTAATGG CCTTATCAA GGCAAAAGGA TTTTGATTTT | 2460 |
| TATAGAAAGG ACTGGCAATA GCAGGATTTT CAGAGTAATA AATCCTATCA TCGCCTTCCA | 2520 |
| AAATCCAGAC CTCTGTTAAT AGGGGATAGT GATTAAAACG GATAGAATAT TCTTTACTAG | 2580 |
| TTTGACCTGT ATGAACCACA AAATTCAGC TTTCTATAAC ATGTGAACCT GGGTGTTCOA | 2640 |
| AGCTAAATAA AGCTCCAAAA TAATCTTCTT TGTAGGTTAG CAAATCAATT CGTTGATCCT | 2700 |
| GACTTTTAC AAAGGAGCAA GTGTCATATT CTCCATTCTT ACGATGGTAA TGAATGCGCA | 2760 |
| TAGGGTAGTT ATACATTTTT TATTTTTCCT TTTTACTTTG TTTCTATTTT ACTAATAAAT | 2820 |
| TTTTGTCAAT CTCGTCTCAA TTAACAGACA TAGTCATATT CTCTAAACTC TGTTTTTAAA | 2880 |
| CGATCCATTA CAAACTTTCT AGCCATGCCT CATCTCTGAC CTGGATACCA AGTTCTTGTC | 2940 |
| CTTTTTGCAG TTTACTTCCA GCGTCTGCAC CTACCACGAC GAGGTCGGTC TTTTGTAGAA | 3000 |
| TACTACCTGT CACTTTGGCA CCCAGACTTT CGAGTTTACT TTTAGCTTCT GAGCGCTTGA | 3060 |
| GTCGTTCCAA TTTTCTGTC AATACCACGG TCAAACCTGA CAAGGCCGCA TCCGCTACTA | 3120 |
| CCGTCTGTCC TTTATAGTCC AGATTGACCC CAGTTTCTTT CAATTCTCTG AGCAGAATTT | 3180 |
| CAGAGCCTTC TGTCGCAAAA TAAGTCTGAA GACTTTTGGC AATCAGCCA CCTAGACTTT | 3240 |
| CAATACTAGC CACTTCCTCT GAATCTGCCT GAGACAGATT TTCAATTGAA TGGAAATATT | 3300 |
| GAAGTAAAG CTGACTAACC TTGCTTCCGA CATGACGAAT TCCCAAACCA AATAAGAGCT | 3360 |
| TCTCGGCAGA ATTTTCCTTT GATGCTTGA TAGCCTGATA CAGTTTAGCA GCGGACTTTT | 3420 |
| CCTTAACTCC CTCTAAAAGG AGGAAATCCT CTTCTTGCAA ACGATAAATA TCCGCCACAT | 3480 |
| CCTTGACTAA ATTAGCAGCA AAAAGCTTCT CAACAATAGA TGGACCAAGG CCTGTAATAT | 3540 |
| TCATAGCATC ACGAGAAGCA AAGTGAATCA AGCCTTCCAT GATTTGAGCA GGGCAACGCG | 3600 |
| GATTGATACA ACGTAGGGCC ACTTCATCTT CAAAGTGCAA CAAGTCAGAG TTACAACCTG | 3660 |
| GACAGTTTGT AGGGATATCT AGTTTTTCTT CAGAAACCG TTTGGACTCT ACCACACGTA | 3720 |
| AAACGGCAGG GATGATGTCA CCAGCCTTAT ATACAATGAC CGTATCGTCT TTTCCGATAT | 3780 |
| CTTTTTCAGC AATATAATCT ACATTGTGCA GGGTCGCACG GCTAACAGTC GTACCGGCAA | 3840 |
| GTTGTACTGG TGTTAGATTA GCAGTTGGAG TTACAACACC GGTACGGCCA ACTGTCCAGT | 3900 |
| CAACTGATAA GAGTTGAGCT TCTTTTCTT CGGCAGGAA CTTGTAGGCT ACTGCCACT | 3960 |
| TTGGAGCCTT AACTGTAAAA CCAAGTTCTT CTGACTTGC TAGGTCGTTG ACCTTGATTA | 4020 |

903

| | |
|---|------|
| CCACTCCATC AATATCGTAA GGCAGATTTT CCCGTTCTTG TCCTACTTCT TGGATAAAAT | 4080 |
| TCCAGATTTC ATCTATGTTT TCAGCCAAGA TTCGCTTAGG ATTGACCACA AAACCTAGTT | 4140 |
| GTTCTAGGTA CTTCAAACCC TTTTCTTGGC TATCACGAGT TGAAGGGCTG GCTTCTTGAT | 4200 |
| AGAGAAACGT TGCAAGATTA CGCTTGGCAA CTA CTGCTGT ATCCA ACTGA CGCAGAGTTC | 4260 |
| CTGCTGCCGC ATTACGAGGA TTAGCAAAT CAGGCTCTCC ATTTTCTTGG CGCGCTTGGT | 4320 |
| TAAGTTGGTC AAAGGAAGCG CGTGGCATGT AACATTCCCC ACGAACTGTG ATATCTAGTT | 4380 |
| CTTCTGGCAA AGTCAAAGGG ATGTCTTAA CACGCTGAG GTTTTCTGTG ATATTTTCAC | 4440 |
| CAATTGAACC ATCTCCACGT GTTACCCAG CAACCAAAT CCCCTTTTCA TAAGTCAGCG | 4500 |
| AGATAGATAA GCCATCGATT TTCAGCTCAC AAATATAGGT CGGATGAGCC ACTTCCTTAC | 4560 |
| GAACACGCGC ATCAAAAGCA TCTAGCTCCT CACATGAAA AGCATCTGC AAATAATAA | 4620 |
| GAGGATACTG ATGACTGTAT TTTTCAAAC CATCTAAAAC CTTGCCACCA ACACGATGAG | 4680 |
| TCGGACTGTC TGCTAGCACT TGCTCTGGAT AAGCAGTTT TAAGTCGACC AACTCACGGT | 4740 |
| AAAGGCGGTC ATACTCACTG TCTGAAACCG AGGATTATC GCTGGTATAG TACTCAGTCG | 4800 |
| CATAGCGATT GAGCAAAGCG ACTAACTCAT TCATTCTTTT ATTCATAAGA CCATTTTACC | 4860 |
| ATAAAACAAG CCCTCCTCAC AAACGAGAAG GCGGAAAAA ACACCTTAGTT TGAATTTATT | 4920 |
| TTTGAAACTC AAGCAACCTT ATATCAATTT TTCAAAATGA GTTCGAACAT ATCCGAGAGC | 4980 |
| TAAGAAATAT AAGGCTACAA CTCCAAGTCC AATAATCAAG AAAGAATAAA GATGGACACT | 5040 |
| TGGCAAGACT GTCATAAATC CTTTGTCAAT AGGCATAAAT AGAATAGCTA AGGTAAAAAT | 5100 |
| TGTACTCAGT ACTCTTCCAA GAAATTCGCT CTCAACCTTG GTTTGTACTT GAGTAAAAAA | 5160 |
| GTGAATATTA AAAATCGTCA TAAACAATTC ACAAATAAA TTTCCAGAAA AGGAAAGAAA | 5220 |
| AGTTGGAAGT GGTAATCCCA TCATAAAAC TCCGACACCT GTCAAAGCCA GTAAATCAA | 5280 |
| AAGATTATAA ATATTAGCTT TAATTTTACT AGCTAGAAGA GCCCAATGA TGGAAACCAAT | 5340 |
| AGCCCCATA GTTAAATAC TTGCATAGGC TCCTTCTGAC CCGTAAAGCT GATTGAAAAA | 5400 |
| GGGAAGTAGA AATTCAAAG CTGCAAAAAA GAAATTAACG CTGGAAGCTA CCAGCAAAAG | 5460 |
| GAAGAAAATT TCTTGCTGAT GCCAGATATA GTGTAACCA TCCTTGATAT CTACAAAAAT | 5520 |
| ATCTCTCCA GTAAAAGCCT TTTTCTCTTG AACTTTTGCT TCCTCTTTTG GAAGGAAAGC | 5580 |
| CACTAGAACA AAAGCAATGA AAAAAGTCAG CGAGTCTAGC AGTAGCGTCA TATGGAGACT | 5640 |
| TGCAAACTGT AAAACAAGGA AGGAAAGAAC AGGAGAGCTA ACACCTACAA CCTGCAAAAC | 5700 |
| CAGCTCTAAG CGAGAATTAT AGATCACAAT CTCATCTTTC TCCACCACTT CAGTTATGAT | 5760 |

904

| | |
|--|------|
| AGCTTTATTG GCTGTGCGAG AAAAGGCAAA AGCAATAGCC TGCACAATGT TAGCAACAAT | 5820 |
| CAAAGCGCCA ATCATCCAGC TATCATTCCT TATGAAAGAA ATAGCCAGAC AAAGAATCCC | 5880 |
| ACAAACAAGA TCTGCCGTCA TTAAATCTT ACGACGAGAA AAACGGTCTG AAATAACTCC | 5940 |
| GCCAAAGGGA TTGACGAGAA TAGATGTGAC GAGCTCAGAA ATCTGATACA TTCCTAAAC | 6000 |
| TGTCTGTCTT ATAGTCCCCA TAGAAGCCAA CCAGACACTA TTCCATAAT CATAGAGCAT | 6060 |
| ATTTCCCATT TTATTGATAG CCCACGGCT AATCAACTGC ACTGCATAGC GATTCATATT | 6120 |
| AAAGCTCTC TCAAATTTTG AACTATTGT ATCAAAACCG AAAGGAGCTT TTTATTTT | 6180 |
| CCCTTATTG GGAAATTAA CTTTGTACAA ATTTTTCGTA GTGTCCTGA TAATAGGCTA | 6240 |
| CTTGCTCTGG AAGACCTAAC ACATCAAAA TATGCATGGC CTCTTGCATC TGCTTACAGC | 6300 |
| CTTCTTTACA CTGTCCTTT TGATATAAGG CAAAACCTT TAAATAATGG AAAACATTAC | 6360 |
| GCTCATAAAG CTTAATACCT TTGTCAATAA TCTTCTCTGT ATAAGCCTCA AAATAGTTGG | 6420 |
| CATTATAAAA AGAAGATGC TCTAAACAAT GCTGGTAACA ATTGAGGGCC AAAATCAACA | 6480 |
| CTAATCTCTT ATGGCGACTA ATCTCTTGGT AAAATTCCTC CCTCTCCATA ACTTCTCTAC | 6540 |
| CAATCCGAGT GACATAGTCT ACATCGTAGA AACTATAGAG GTTACCGAAA AGAATCAACT | 6600 |
| CATACATGGT CCATTCTTCT GTTTTGAAGA GATAATCTGC TACCTTACCC AAATCATCCT | 6660 |
| GCTTCATATC ATAACCTGCA TCTCTTTGAC AAATCAGACC TTGTAGCAAA ATCCAGTTCA | 6720 |
| GCTCAAAATA AAGGGGAGTC GTCGAACCTT TAGACTTTTC AAGTTGTTCT CTTTGAAGCT | 6780 |
| TTTGAAAACC TGCAATATCG TTTGAATAGT AAAGTGGGAT AATCTGTGCC ATCATAGACA | 6840 |
| CATGTTTCATG ATTATGAAAA TTCCTTGCCT TATCCATGAA ATTTTCGATT GTTACATGAA | 6900 |
| TGTTATCCAA AATCTCAAAG AAACGGGAGA CTGCCAGGTC AGACTCCCCA AGCTCAAAGC | 6960 |
| GAGATAACTG AGAGGTAGAG CAGGATTCGC CTGCTGCTTC CTTTAAAGAA TAATTTCCAC | 7020 |
| TTGTTGCAAA TTCACGAAAT ACTTTTCCAA GATGTTCCAT CTTTACACCT GCTCTGATAA | 7080 |
| TTCTTCCCAC TCAAGCATAG CTCTTCCCTG ACGATGGCTG ATTTTGTCCA GCTCAGCCTG | 7140 |
| TAATTCATG AGTTTGTGCG CATCGTTTGT TTCCAACATT TGTTTCAGAA TGGCTTGGCT | 7200 |
| TTGACTTTCT AGCTCTTCAA TTTCAGCTTC TAGACTTTCG ATTTGTGCA TGAGTTTTCG | 7260 |
| AACTTCTTTT TGACTTTCTT TCTGGGCCTG ATAGTCATG ACTGGACTTG CTTCCCTTGC | 7320 |
| TTGATTGCTA GTTGAAGCTT CCTCAGTCTG ACTCATTTCT GCTGTTGCTT TCTTCTCAAC | 7380 |
| ATAGTAGTCG TAATCTCCAA GGTAGAGAGT TGAACCATTC TCAGACAATT CCAAACATG | 7440 |
| AGTTGCCACA CGATTGATAA AGTAACGATC ATGACTGACA AACAGCAAGG TTCCATCAAA | 7500 |
| GTCAATCAAG GCATTTTCTA GCACTTCCTT ACTATCAATA TCCAAGTGGT TGGTCGGCTC | 7560 |

905

| | |
|--|------|
| ATCCAGAATC AAAAAGTTAT TGTTTTCCAT AGACAATTTA GCTAAAAGCA AACGAGCTTT | 7620 |
| TTCGCCACCA GATAGCATGC CGACTGATTT TTTAACATCA TCTCCTGAGA AAAGGAAGGC | 7680 |
| TCCAAGACGG TTGCGGATTT CAACTTCTGG TGTCAGTTTG AAATCATTCC AGAGTTCATC | 7740 |
| CAGCACCGTA TTA CTGGTG TCAGCTTGCT TTGGGTTTGG TCATAGTAAC CAACCTCAAC | 7800 |
| ATTAGCGCCA AAGCGCTTTT CTCCCTTGAT AAAAGGAATC TGGTCCACAA TAGACTTGAT | 7860 |
| AAAGGTTGAC TTGCCGATAC CATTGAGACC AACGATAGCG ACAGCATTCA TCTTACGAAG | 7920 |
| ATCTAGGTTA ATCGGTTGTG ACAAGACTTC CCCGTCATAG CCAACAGCTG CATTTTCAAC | 7980 |
| AGTCAAAACA ACATTGCCCCG ACGTTTTTTC AGACTGGAAG GTCATGTTGG CTGATTTCTT | 8040 |
| GCCAGCTTCA GGCTTGTCCT AACGTTCCAT TTTTCCAGT TGTTTACGGC GAGATTGAGC | 8100 |
| ACGTTTAGTC GTTGAAGCAC GAACTAGATT GCGATTGACA AAGTCTTCCA GAGCAGCGAT | 8160 |
| TTCCTTCTGT TGCTTTTCAT AGTTTTTTCG CTCAGTAACT AGCTTTTGCT CCTTCAATTC | 8220 |
| GACAAAACGA GAGTAATTCC CCACATAGCG ATCCAAGGAA TGCTTGGTCA AATCTAGCGT | 8280 |
| AATTGTCGCA ACCTTGTCCT AGAAATAACG GTCGTGGCTG ACGATAATGA GGGCACCCT | 8340 |
| ATAGTTTACC AAGTAATTCT CTAGCCAGGC GATGGTTTCA ATATCCAAGT GGTTAGTTGG | 8400 |
| CTCGTCCAAG ACCAAGAGAT TGGGCTTTTC AAGGAGCATT TTGGCAAGTG CCAAACGAGT | 8460 |
| ATTTTGACCA CCAGAAAGCT CAGCAATTTT CATCTGCCAC ATAGACTCGT CAAACTTGAA | 8520 |
| TCCATTCAAA ATCGCTCGAA TATCAGCTTC ATAGGTAAAG CCACCTGCTT GGCGAAAATT | 8580 |
| CTCAGATAAG CGGTCATAAT CTGACATCAG TTTATCCAAA TCCTCACCAG ACTTTTCACC | 8640 |
| CATCTCCAGC TCCATCTGAC GCAGTTGTCT CTCCGTCCGA CGCAAATCAT TAAAGACATG | 8700 |
| AAGCATTTCA TCGTAGATGG TATTTTCAGA CTCAAAACGG CTATCTTGGG CTAGGTAAGA | 8760 |
| CAGAGAAATA TCTTTTCTCT TATTGATTTT TCCGCTAGTT GGCTCCTCTT CTCCAATAA | 8820 |
| AATCTTCAAA AGAGTAGACT TACCTGCACC ATTTTTCCTA ACAAGAGCAA TCCGATCTCG | 8880 |
| TTCATCAACC TGCAGGTTGA TATTATCGAA AAGAACCCTCT CCTGCAAAAG AACGTTCAAT | 8940 |
| TTTATTAGCT TGTAATAATA TCATACAAGT AGTATAGCAT GTTTCCTTAA GGCATTCAAG | 9000 |
| ATAATCGTAA GTCTTTTAGT ACAACTTTTA TAACATAAAA TAACTAAAT TATGTATATT | 9060 |
| TTATATTAGA TTA CTCACT ATCTTGTTGG ATTTTCTAAC CAGCTAATCT TGTTTCAAAT | 9120 |
| AGTTATCGCA CAAGTCTATT ATTTAATTCT TTTCATCATT TACGTACGTA TAGCAGATTG | 9180 |
| AAATAAGATG AGAACAAATC GATTGGGAAA GTAAAATTAA TTTCTATAAA TGTTTTAGCA | 9240 |
| ATTGTTTCGT ACTATTTTAG ATTCAGTCTA CTATATACAA TATTTTCGGA ACATTCAACT | 9300 |

906

| | |
|--|------|
| TTTTAACTCT ATTTATTACT AGATTTCATA ATTAAAAAAC CTA CTGACCA AGCTAGAAAG | 9360 |
| CTTGATACAA TAGGCTTTT AAAGACTGAT TATTTAACAG CGTCTTTAAG AGCTTTACCA | 9420 |
| GCTTTGAATG CTGCTACTTT AGAAGCTGCA ATGTGCATTT CTTTACCAGT TTGTGGGTTG | 9480 |
| CGACCTTTAC GTTCTGCGCG CTCACGAACT TCAAAGTTAC CAAAACCGAT CAATTGAACT | 9540 |
| T | 9541 |

(2) INFORMATION FOR SEQ ID NO: 133:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3502 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 133:

| | |
|--|------|
| TTGACTATCC TATCATGCTT TCTAAGGTCT ACTCAAGAA ATCATTTTCA AGTTTTCACA | 60 |
| CCTTTCTCAA AAAAGTTAAA AAATTTTCTC AAAAACGCTT GACTCTGACC TAAGGCGAAG | 120 |
| GGTTATACTA TCATTGTAAG GAGGAAATCA TGTACCATAT AAAAGAAGCT GCGCAGCTTT | 180 |
| CAGGTGTCTC TGTCAAGACC CTGCATCACT ATGACAAGAT AGGACTCTTG GTCCCTTAA | 240 |
| AGTCGAAAA CGGCTATCGA ACCTACAGTC AAGAGGATTT GGAACGCCCT CAGGTCATTC | 300 |
| TTTACTACAA ATATCTAGGC TTTTCTTTAG AGAAAAATAGC AGAGCTGTTA AAGGAAGAAA | 360 |
| GGACAGATTT ATTGCCCCAT TTGACTAGGC AGTTGGACTA TCTAACTCGC GAAAGGCAAC | 420 |
| ATCTGGATAC CTTGATTTCC ACCTTGCAAA AACTATTCA AGAACAAAA GGAGAAAGAA | 480 |
| AAATGACCAT TGAGGAAAAA TTCACGGGAT TTAGCTATCA AGACAATCAA AAATACCACC | 540 |
| AAGAAGCGGT AGAGAAATAT GGTCAAGAAG TCATGGGACA AGCGCTCGAA CGCCAAAAAG | 600 |
| GTCACGAAGA CGAGGCTACG GCCGCCCTTA ACCAAGTCTT TCAAACCTTG GCACAAAATC | 660 |
| TTCAAGTTGG TTTACCTGCA ACAGCAACCG AAAACCAGGA GCAAGCAGCC AAGCTCTTGC | 720 |
| AAGCCATTCG CACTTATGGA TTTGACTGCT CTATTGAGGT ATTCGGTCAT ATCGGTAAAG | 780 |
| GTTACGTCTA CAACCCAGAG TTTAAGGAAA ACATTGACAA GTTTGGTTCT GAAACAGCCC | 840 |
| AGTACACGTC AGATGCCATT GCGGTTTACG TTCAGACAAA TGCAGAATAA ATAGGCTAGG | 900 |
| AATTTCTTAG CCTATTTTTT ACTTCAAATC ATAAAGCCAG TCGTCACCGT TTTTGTAAGTA | 960 |
| AAAGAATCA CTGAGATCTT CTTCTAGAAA CACACGAAGC ATATCAGACA TATCATCGGT | 1020 |
| TGCAAGTTTT AGATGAGAAA GATTTTCAAA GTCCTCCAC CAAACTTTCC CTTCTGCTGA | 1080 |
| AGACTGGAGT TCACCAGTAA AGTGTCTGT CTTGTAAAAA AGGACGACAT AACGATAATC | 1140 |

907

| | |
|---|------|
| CTTGTCGTCA TACCAGTTT TGATACCACA GAGTTGGGGT TTGAAAATGA TCAGACCAGT | 1200 |
| TTCTTCTTTC ACTTCACGAA TGACAGCATC GACAAAGGAT TCGCCACGTT CAACATGACC | 1260 |
| ACCAGGAAAA GTAATGCCAG ACCAGTCGGG ATTAACTCGG TCTTGGACCA GGACCTTATC | 1320 |
| TCCGTTTTTA ATCATACACA TGTTAACAAA TTCGACTGCC TCTCTTCTGT TCATTCTTCA | 1380 |
| CAACCTTTAA TCTTTAATCA TAATGCAGAC TTCCCGCCAC CCAGCCGGTA CAGAGGGCAG | 1440 |
| AAGTGATGTT AAAGCCACCC GTGTGGGCAT TGATATCCAT AACTTCGCCT GCAAAGTGGA | 1500 |
| GGCCAGGTAC CAGCTTACTT TCAAGGGTTT TAGGATTGAT TTCCTTGAGA CTGACTCCAC | 1560 |
| CCTTGGTAAC AAAGGACTTT GCAAGGGACA TTTTTCAGT TACAGGAATT TTAAGTTCTT | 1620 |
| TAATGGACTG GACAAGTTGT TCTCGTTCCT TTTCAGTCAG TTGTTTGA CTGACTTTCAGGAT | 1680 |
| ATCCTTGTAC AAAAAATCG GCCAAGCGT CTGGTAACAA GGTTTTTAAA GCGTTTTTCA | 1740 |
| AGGATTTTTT CCGATTTTCT TCTAGAAATG TAACCAAGTC CTTCTCAGAA AGTTGAGGCA | 1800 |
| AAACATCGAG TGAGAGAACC TCCCCACCTT TGACAAAGCT AGACATGCGT AGGGCAGCAG | 1860 |
| GACCTGACAA ACCAAAGTGG GTAAAGAGTA AATCATGAGT GATGACATGC TTACCATAAC | 1920 |
| TTAGGGTCAC ATCGTCCAGA GAAATACCTT GTAAGGCTTT ATGTGGAAAA TCTGTTAATA | 1980 |
| AAGGACTTTC AGCAGCCTCA AGATCGGTGA TGGTATGCTT AAAATGGCGA GCAATCTCGT | 2040 |
| GACCAAAACC AGTCGAACCA GTCGAAGGAT AAGACTTACC ACCTGTTGTG ACAATGAGTT | 2100 |
| TCTCACAAGT GAAGGTTTGA TCCGCTGACT TAAGGACAAA CTGGTCATCT ACTTTTTTAA | 2160 |
| CAGAAACGAT TTCTATTGA GTAGCAACTT GACCACCTAG TTCGGTGATT TTCTTTTCCA | 2220 |
| AAGCTTCGAT AATAGTCCGA GACTTGTCAC TGGCTGGAAA GACGCGTCCG TGGTCTTCGA | 2280 |
| CCTTAAGTTT AACACCATT TCTGTAAAA AGTTGATGAT GTCATGATTA TCGAACTGGG | 2340 |
| AGAAAACACT GTAAAGAAAG CGTCCGTTTC CAGGAATTCC AGCTAGCAGG TTGTCTAAGC | 2400 |
| TACCATTGTT GGTCACATTG CAACGTCCCC CACCAGTCCC AGCTAATTTT TTTCCAAGTT | 2460 |
| TCCGATTTTT TTCGATGAGG AGGGTTTTCT GTCCATAAAA GCTACTGGAA ATCGTAGCCA | 2520 |
| TCATACCAGC AGGTCCCCCA CCGATGACAA TAGTATCAAA ATGTTTCATA GCTCTATTGT | 2580 |
| ACCACAAAAA AACAAAGAT GATGGTCACC TCTTGTCAG AATGCAATTA ATCAATTTCA | 2640 |
| TAGCCCATCA GCAAACCGCC CTCTTCTGCA TAGAACTGC AGAGACCAGA GGTGGTAGA | 2700 |
| ATTTTAATAT CCGCTTGTGG GAAGGTTTCA CGGATTCGCT CTGAGAGCTG TTGACAACAT | 2760 |
| TTTTCGTTAT TGCGTTGGGC CATGACAATA CGGCCACCAG CATATCCAGC TTTTACTAAC | 2820 |
| TCATCATAGG CAGCTTGAAC TGATTTCTTT GATCCCCTTG CTTTTGTAG CAATTCGAGA | 2880 |

908

| | |
|---|------|
| GTCCCAGTTT CACTAGCTTT TCCGACCATA CGAATGTTGA GAAGGCCAAC GACCGTACCG | 2940 |
| ATAAGCTTGC TCAAACGGCC GTTCTTCACC AAGTTATCGA CTTTGGCTAG GACAAAGAGC | 3000 |
| AACCTAGTTT TTTCTTGATA GGCGGTGATA GCTTCAACCA CTTCTTCAAA AGACAAGCCC | 3060 |
| TGGTCAATCA AGTCATTCAA TTTTCTACG AGTAGGTCAA CTTACCACC AGCAGATAAA | 3120 |
| CTATCAATCA CATGAATCTT AGTGTACGGA TGGTCTTCCA GATAAATATT CTTTGCTAGT | 3180 |
| TGAGCACTAT TGTGACTGCC AGAAAGGGTA CCTGTGATGG TTAGTAGGAA AATGTTTTTG | 3240 |
| GCACCTTCAA ATGCTCGCAA ATAGTCATCT GGGCTTGGAC AAGCCGATTT TGAAGCTTCT | 3300 |
| GCAGTTGCAT ACATGGTTTC CATCATTTGG TCAATATCGA GACTGGCGTC ATCAACAAAG | 3360 |
| ACCTGATCAG CTAATTGAAT GGTAAAGGG ACACCTACAA AGGTTGTGTT AATAGCTGGT | 3420 |
| GTTGGCAGTT GACGATAATC ACAACCAGAG TCAGCAATAA TCTTCCAAGT CATAGAAATT | 3480 |
| CTCCATCTTT GTCAGGAACG AT | 3502 |

(2) INFORMATION FOR SEQ ID NO: 134:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 12665 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 134:

| | |
|--|-----|
| CGATTGATTT TTTTAAAGCG TTCGATAGAG AATGAGAAAC GAATCCTTAG CAATGGCGGG | 60 |
| AAAGAATTG GAGTTGAGAA TACAAAACGA TTAATATGCTCATATTGT TTTTATCTC | 120 |
| TCTTGCTTGG TTGAGGCAAT GGTGCACAAG ACAATTTTGT ATGGCATGGG CATGGTTGGT | 180 |
| TTAGTCTTGC TTATTTTTC TATGCTGATG TTGATGTTGG TGATTCACCT GTTGGGAGAT | 240 |
| ATTTGGACAG TGAAGCTTAT GCTTGTCAAT AATCACAAT ATGTAGATCA TATCTTGT | 300 |
| AGGACAGTAA AACACCCTAA TTACTTTTTA AATATCTTCTC CTGAGTTGAT TGGCTTGACC | 360 |
| TTGTTGAGTC ATGCTTATGT GACTTTTGT TTAGTTTTC CAGTTTATGC AGTTATTTG | 420 |
| TATCGACGAA TAGCTGAAGA GGAAAAGCTA TTACATGAAG TTATAATCCC AAATGGAAGC | 480 |
| ATAAAGAGAT AAATACAAA TTCGATTAT ATACAGTTCA TATTGAAGTG ATATAGTAAG | 540 |
| GTTAAAGAAA AAATATAGAA GGAAATAAAC ATGTTTGCAT CAAAAAGCGA AAGAAAAGTA | 600 |
| CATTATTCAA TTCGTAAATT TAGTGTGGA GTAGCTAGTG TAGTTGTGTC CAGTCTTGT | 660 |
| ATGGGAAGTG TGGTTCATGC GACAGAGAAC GAGGGAGCTA CCCAAGTACC CACTCTTCT | 720 |
| AATAGGGCAA ATGAAAGTCA GGCAGACAA GGAGAACAAC CTAAAAAAT CGATTACAGAA | 780 |

909

| | |
|--|------|
| CGAGATAAGG CAAGGAAAGA GGTCTGAGGAA TATGTAAAAA AAATAGTGGG TGAGAGCTAT | 840 |
| GCAAAATCAA CTAAAAAGCG ACATACAATT ACTGTAGCTC TAGTTAACGA GTTGAACAAC | 900 |
| ATTAAGAACG AGTATTTGAA TAAAAATAGTT GAATCAACCT CAGAAAGCCA ACTACAGATA | 960 |
| CTGATGATGG AGAGTCGATC AAAAGTAGAT GAAGCTGTGT CTAAGTTTGA AAAGGACTCA | 1020 |
| TCTTCTTCGT CAAGTTCAGA CTCTTCCACT AAACCGGAAG CTTCAGATAC AGCGAAGCCA | 1080 |
| AACAAGCCGA CAGAACCAGG AGAAAAGGTA GCAGAAGCTA AGAAGAAGGT TGAAGAAGCT | 1140 |
| GAGAAAAAAG CCAAGGATCA AAAAGAAGAA GATCGTCGTA ACTACCCAAC CATTACTTAC | 1200 |
| AAAACGCTTG AACTTGAAAT TGCTGAGTCC GATGTGGAAG TTAATAAAGC GGAGCTTGAA | 1260 |
| CTAGTAAAAG TGAAAGCTAA CGAACCTCGA GACGAGCAAA AAATTAAGCA AGCAGAAGCG | 1320 |
| GAAGTTGAGA GTAAACAAGC TGAGGCTACA AGGTAAAAA AAATCAAGAC AGATCGTGAA | 1380 |
| GAAGCAGAAG AAGAAGCTAA ACGAAGAGCA GATGCTAAAG AGCAAGGTAA ACCAAAGGGG | 1440 |
| CGGGCAAAAC GAGGAGTTCC TGGAGAGCTA GCAACACCTG ATAAAAAGA AAATGATGCG | 1500 |
| AAGTCTTCAG ATTCTAGCGT AGGTGAAGAA ACTCTTCCAA GCCCATCCCT GAAACCAGAA | 1560 |
| AAAAAGGTAG CAGAAGCTGA GAAGAAGGTT GAAGAAGCTA AGAAAAAAGC CGAGGATCAA | 1620 |
| AAAGAAGAAG ATCGCCGTAA CTACCCAACC AATACTTACA AAACGCTTGA ACTTGAAATT | 1680 |
| GCTGAGTCCG ATGTGGAAGT TAAAAAGCG GAGCTTGAAC TAGTAAAAA GGAAGCTAAG | 1740 |
| GAACCTCGAA ACGAGGAAAA AGTTAAGCAA GCAAAAGCGG AAGTTGAGAG TAAAAAGCT | 1800 |
| GAGGTACAA GGTTAGAAAA AATCAAGACA GATCGTAAAA AAGCAGAAGA AGAAGCTAAA | 1860 |
| CGAAAAGCAG CAGAAGAAGA TAAAGTTAAA GAAAAACCAG CTGAACAACC ACAACCAGCG | 1920 |
| CCGGCTCCAA AAGCAGAAAA ACCAGCTCCA GCTCCAAAAC CAGAGAATCC AGCTGAACAA | 1980 |
| CCAAAAGCAG AAAAACCAGC TGATCAACAA GCTGAAGAAG ACTATGCTCG TAGATCAGAA | 2040 |
| GAAGAATATA ATCGCTTGAC TCAACAGCAA CCGCCAAAAA CTGAAAAACC AGCACAACCA | 2100 |
| TCTACTCCAA AAACAGGCTG GAAACAAGAA AACGGTATGT GGTACTTCTA CAATACTGAT | 2160 |
| GGTTCAATGG CGACAGGATG GCTCCAAAAC AATGGCTCAT GGTACTACCT CAACAGCAAT | 2220 |
| GGCGCTATGG CGACAGGATG GCTCCAAAAC AATGGTTCAT GGTACTATCT AAACGCTAAT | 2280 |
| GGTTCAATGG CAACAGGATG GCTCCAAAAC AATGGTTCAT GGTACTACCT AAACGCTAAT | 2340 |
| GGTTCAATGG CGACAGGATG GCTCCAATAC AATGGCTCAT GGTACTACCT AAACGCTAAT | 2400 |
| GGTTCAATGG CGACAGGATG GCTCCAATAC AATGGCTCAT GGTACTACCT AAACGCTAAT | 2460 |
| GGTGATATGG CGACAGGTTG GGTGAAAGAT GGAGATACCT GGTACTATCT TGAAGCATCA | 2520 |

910

| | | | | | | |
|------------|------------|------------|------------|-------------|-------------|------|
| GGTGCTATGA | AAGCAAGCCA | ATGGTTCAAA | GTATCAGATA | AATGGTACTA | TGTCAATGGC | 2580 |
| TCAGGTGCCC | TTGCAGTCAA | CACAACTGTA | GATGGCTATG | GAGTCAATGC | CAATGGTGAA | 2640 |
| TGGGTAAACT | AAACCTAATA | TAAGTAGTTA | ATACTGACTT | CCTGTAAGAA | CTCTTTAAAG | 2700 |
| TATTCCCTAC | AAATACCATA | TCCTTTCAGT | AGATAATATA | CCCTTGTAGG | AAGTTTAGAT | 2760 |
| TAAAAAATAA | CTCTGTAATC | TCTAGCCGGA | TTTATAGCGC | TAGAGACTAC | GGAGTTTTF | 2820 |
| TGATGAGGAA | AGAATGGCGG | CATTCAAGAG | GCTCTTTAAG | AGAGTTACGG | GTTTAAACT | 2880 |
| ATTAAGCCTT | CTCCAATTGC | AAGAGGGTTT | CAATCTCTGC | CAGGGTGCTG | GCTTGCGAAA | 2940 |
| TGGCTCCACG | GAGTTTGGCA | GCGCCAGATG | TTCCACGGAG | ATAGTGAGGA | GCGAGACCGC | 3000 |
| GGAATTCACG | AACTGCGACG | TTTCTCTCCT | TGAGGTTAAT | CAATCGTTC | AAGTGTTCGT | 3060 |
| AGGCGATCTT | CATCTTGCT | TCAAAGGTCA | AATCAGGTAG | GATTTCTCCT | GTTTCAAAGT | 3120 |
| AATGGTTGAT | TTGGTTGAAG | AGGTAAGGAT | TTCCCATGGC | AGCTCGGCCA | ATCATGACTG | 3180 |
| CGTCAGCACC | AACTTCTTCG | ATGCGTTGCT | TGGCTTCTTG | GACAGTACGG | ATATCACCGT | 3240 |
| TGGCGATGAA | TGGAATCTTG | GTTAGAGCTT | GGGCAACCTT | GTAAAGGGTC | TCAAGGTCTG | 3300 |
| CGTGGCCAGT | ATACATTTGT | TCACGGGTAC | GGCCATGCAT | GGCGAGGGCA | GAAACACCTG | 3360 |
| CAGCTTCAGC | AGCGAGAGCA | TTTTCTACTG | CAAGAGATGG | GTCCGCCAG | CCGGTACGCA | 3420 |
| TTTGTACAGT | AAGTGGGATA | TCAAGGACAG | ACTGGACCTT | GTTGATGATG | GAGTAAATCT | 3480 |
| TGTCTGGATC | CTTGAGCCAC | ATAGCACCAG | CTTCGTTCTT | CACGATTTTG | TTGACAGGGC | 3540 |
| AGCCCATGTT | GATATCGACG | ATATCGGTCT | TGGTGTTC | TTGGATGAAT | TCTGTGCGC | 3600 |
| GTGCTAGGCT | GTCTTCATCG | CTACCAAAAA | GTTGGATAGA | GACAGGGTTT | TCGCCCTCAT | 3660 |
| CGATATGAAG | CATGTGCAGG | GTTCCTTCGT | TGTTGTATTG | GATTCCCTTG | TCAGAGACCA | 3720 |
| TTTCCATTAC | AACGAGTCCA | GCTCCGAGCT | CCTTTGCGAT | AGTACGAAAG | GCTGAGTTGG | 3780 |
| TCACGCCAGC | CATAGGCGCT | AAAACGGTAC | GATTGGGAAT | CTCAATATTG | CCAATCATAA | 3840 |
| AAGGTGTATT | AAGATTGTG | ACGAATGAGT | TCCTCCAGGT | CCTTTTCATC | AAAGTTGTAA | 3900 |
| GTAGTTTGGC | AGAATTGACA | AGTGATTCT | GCCCCGTGGT | CTTCCTCTTT | CATTTCCCTGT | 3960 |
| AAGTCTGAGC | TTGGAAGGCT | GGCAAGAGCG | TTCATAAAGC | GTTTCATGGCT | ACAGTCACAT | 4020 |
| TGGAAACGGA | TTTCTTCTTC | AGAAAGACGC | TTGTAGGCTT | CGTCCCGTA | GATAGCCTTG | 4080 |
| AGGAGGGCTT | CGATATGGTC | GTCGCTTTCG | AGAAGAGTAG | AGATAGCTGG | CATTTCTTGG | 4140 |
| ATGCGTTTTT | CAAAGCGAGC | AATCTCTTCT | TTCTTGCTC | CTGGCAAGAC | TTGAACTAGG | 4200 |
| AAACCACCTG | CAACCTTGAC | CTTGCTTCC | TCGTCCAAAA | GGACATTGAG | GCCGACCGCT | 4260 |
| GAAGGCGTTT | GTTGGCTTTC | AGTAAGGTAA | AAGGCAAGGT | CTTCACCGAT | TTCTCCAGAG | 4320 |

911

| | |
|---|------|
| ATGAGGGGAG TTATAGAGTT GTAAGGATTT CCAGTACCGT AGTCTGTGAT AACGAGGAAT | 4380 |
| TGACCATTTC CAACAAAAGG TCCGACTAGG ACTTCACCAG TCGCAGTCTT TTTGATGTCA | 4440 |
| ACACCAGGAT TTTGAACATA GCCTTTGACG TTCCCCCTGG TATCAGCGAC GGTGATAATA | 4500 |
| GCACCTAGAG AGCTAGATCC CAACACCTTA ACTGTAAGTT TGGTATTTCC TTTTTCATTG | 4560 |
| GCTGCGAGAA TCTGGCTAGC GATAAGAGTT CGACCAAGCG CTACAGTTGA GCTAGCTTGG | 4620 |
| GTTTGATGTT TTTCTTGAGC AGTGCGGACG GTTTCAGTGC TATCAAGGAC AAAAGCACGA | 4680 |
| AAGGcTCCGC TTTCTGATAT AGTTTAAATA ATTTTATCCA TAGCTACTAT TTTAGCATAA | 4740 |
| AAATGCCCCA AGGGGGAGCC GTGTGTTTAC TGATTTTCAG GATAATGGAC CAGGAAATCA | 4800 |
| GCATGAAAAT AAAAAGAGAA ACAGATTATT TTAGCATTTG TCAGATTTAT GCTATGCTTA | 4860 |
| AGGTAGAAAA TGAAAGGGAT AACAAATGTA TTTAGGAGAT TTGATGGAGA AAGCCGAGTG | 4920 |
| TGGTCAATTT TCAATACTTT CCTTCTATT ACAAGAGTCT CAGACGACCG TCAAGGCTGT | 4980 |
| AATGGAAGAA ACAGGATTTT CAAAAGCAAC CCTAACCAA TATGTCACCC TGCTCAATGA | 5040 |
| CAAGGCTTTG GATAGTGGCT TAGAGCTGGC TATTCACTCA GAAGATGAAA ATCTGCGTCT | 5100 |
| GTCTATCGGT GCAGCTACCA AGGGGAGAGA TATTCGGAGC TTGTTTTTGG AGAGTGCTGT | 5160 |
| TAAATACCAG ATTTTGGTTT ATCTTCTCTA CCACCAACAG TTTTtagccc ATCAGCTGGC | 5220 |
| TCAAGAATTG GTGATTAGCG AGGCTACGCT TGGTCGTCAC TTGGCTGGTT TAAATCAGAT | 5280 |
| TTTGTCAGAA TTTGATTTAT CCATCCAAA TGGCCGTTGG CGAGGTCCAG AGCATCAGAT | 5340 |
| TCACTATTTT TATTTCTGTC TTTTCCGAAA GGTCTGGTCG AGTCAGGAAT GGAAGGTCA | 5400 |
| CATGCAGAAA CCAGAGAGAA AACAGGAGAT TGCCAATTTA GAGGAAATCT GCGGTGCAAG | 5460 |
| TTTGTCTGCG GGGCAGAAAT TGGACTGGT TCTCTGGGCT CACATCAGTC AACACGTCT | 5520 |
| TCGGGTCAAT GCTTGTCAGT TTCAAGTCAT AGAAGAGAAA ATGCGAGGGT ATTTTGACAA | 5580 |
| TATCTTTTAT CTTCGTTTGC TGAGAAAGGT TCCGTCCTTT TTTGCTGGGC AACATATTCC | 5640 |
| ACTAGGAGTT GAGGATGGTG AGATGATGAT ATTCTTCTCT TTTCTCCTAT CTCATCGCAT | 5700 |
| TCTTCCTCTT CATACTATGG AGTATATTCT TGGTTTTGGA GGGCAGTTGG CAGATTTACT | 5760 |
| GACGCAATTG ATTCAAGAAA TGAAGAAGGA GGAATATTG GGGGATTATA CAGAGGACCA | 5820 |
| TGTCACCTAT GAACTCAGTC AGCTTTGTGC TCAAGTCTAT CTCTATAAGG GCTATATTTT | 5880 |
| ACAGGATCGC TACAAGTACC AGTTAGAGAA TCGTCATCCA TATTTACTGA TGGAACATGA | 5940 |
| TTTTAAAGAG ACAGCAGAGG AGATTTTCA TGCTCTACCT GCTTTTCAAC AGGGGACAGA | 6000 |
| TTTAGATAAG AAGATTCTCT GGAATGGCT CCAGTTAATC GAATATATGG CTGAAAACGG | 6060 |

912

| | |
|--|------|
| TGGCCAGCAT ATGCGGATTG GTCTGGATTT GACATCTGGT TTTCTTGTCT TTTCAAGGAT | 6120 |
| GGCAGCCATT TTGAAACGGT ATTTGGAATA CAATCGTTTT ATTACCATTTG AAGCTTATGA | 6180 |
| CCCTAGTCGG CATTATGATT TGCTGGTTAC CAATAACCCG ATTCATAAGA AGGAACAGAC | 6240 |
| ACCAGTCTAT TATTTAAAAA ATGACTTGGA TATGGAGGAT TTGGTAGCGA TTCGCCAGTT | 6300 |
| ATTATTCACT TAAAAGGCTT GGTTAATCCA GGTCTTTTTT GTGAAATTCA CACAATCTCC | 6360 |
| TCACATTTTT TTAATAATTA AAAAAAGTTG ATAAACAAGA AAGCGCTTTA TTTTGTATAC | 6420 |
| TAGTAAGTGT AAAGAGGAAA CACCTCAAGA TCTTTATCAG GAGGACAGTA CATGTCACAA | 6480 |
| GAAAAATACA TCATGGCCAT TGACCAGGGA ACTACAAGTT CTCGTGCCAT CATTTTCAAC | 6540 |
| AAAAAAGGGG AAAAGGTTAG CTCGAGTCAA AAAGAGTTTA CCCAGATTTT CCCTCAGGCA | 6600 |
| GGTTGGGTTG AGCACAATGC CAATGAAATT TGGAAGTCTG TTCAGTCAGT TATTGCGGGT | 6660 |
| GCTTTCATCG AAAGTGGTGT CAAGCCAAAT CAAATCGAGG CAATCGGGAT TACCAACCAA | 6720 |
| CGTGAAACAA CGGTGTCTG GGATAAGAAA ACAGGACTTC CTATCTACAA TGCTATCGTT | 6780 |
| TGGCAGTCAC GCCAGACAGC ACCTTTGGCT GAGCAACTAA AAAGCCAAGG TTATGTGGAA | 6840 |
| AAATCCCATG AAAAGACTGG TTGATTATT GATGCTTACT TCTCTGCTAC CAAGGTTCGT | 6900 |
| TGGATTTTGG ATCATGTAGA AGGTGCTCAA GAGCGAGCAG AAAAAGGGGA ATTGCTCTTT | 6960 |
| GGTACTATCG ATACTTGGTT GGTTTGAAA TTGACTGACG GTGCGGCTCA CGTGAAGTAC | 7020 |
| TACTCAAATG CAGCTCGTAC CATGCTTTAT AACATTAAAG AACTCAAATG GGATGATGAG | 7080 |
| ATTTTGAAA TCCTTAACAT TCCGAAGCT ATACTTCCAG AAGTTCGTTT TAACTCCGAA | 7140 |
| ATCTACGGCA AGACAGCTCC ATTCCATTTT TACGGTGGAG AGGTGCCAAT CTCAGGTATG | 7200 |
| GCTGGGGACC AACAGCAGC CCTCTTTGGA CAGTTGGCTT TTGAGCCAGG TATGGTTAAG | 7260 |
| AATACTTATG GAACAGGCTC TTTCATCATC ATGAATACTG GGGAAGAGAT GCAGTTGTCT | 7320 |
| GAAAACAACC TCTTGACAAC CATTGGTTAC GGAATCAACG GTAAGGTTTA TTATGCCTTG | 7380 |
| GAAGGTTCTA TCTTCATCGC AGGAAGTGCT ATTCAGTGGC TTCGTGACGG TCTTCGCATG | 7440 |
| GTTGAAAATT CACCAGAATC TGAAAAATAC GTCGTGATT CTCACAACAA CGATGAAGTT | 7500 |
| TATGTCGTTT CAGCCTTTAC AGGTCTAGGC GCTCCATACT GGAACCAAAA TGCTCGTGGT | 7560 |
| TCCGTCTTTG GTTTGACTCG TGGAACAAGC AAAGAAGACT TTATCAAGGC GACTTTGCAA | 7620 |
| TCTATTGCTT ATCAAGTGCG TGATATCATC GACACCATGC AAGTGGATAC TCAGACCGCC | 7680 |
| ATTCAAGTAC TGAAGGTGGA TGGTGGTGCA GCCATGAACA ACTTCCTCAT GCAGTTCCAG | 7740 |
| GCGGATATTT TAGGCATTGA CATTGCACGT GCTAAAAACC TGGAACAAC AGCTCTAGGA | 7800 |
| GCGGCCTTCC TAGCAGGTTT GTCAGTAGGG TACTGGAAAG ACTTGACGTA GTTGAAACTC | 7860 |

913

| | |
|--|------|
| TTGAACGAGA CAGGAGAACT CTTTGAGCCA TCTATGAACG AATCTCGCAA GGAACAACTC | 7920 |
| TACAAGGGCT GGAAGAAGGC TGTGAAAGCA ACTCAAGTCT TTGCGGAAGT AGACGACTAA | 7980 |
| TACTGGCAGA ATAAAGCGAT TTATTTAGAA AGTGTGTAAA TATGGAATTT TCAAAGAAAA | 8040 |
| CACGTGAATT GTCAATTAAA AAAATGCAGG AACGTACCCT GGACCTCTTG ATTATCGGTG | 8100 |
| GAGGAATCAC AGGAGCTGGT GTAGCCTTGC AGGCGGCAGC TAGCGGTCTT GAGACTGGTT | 8160 |
| TGATTGAAAT GCAAGACTTT GCAGAAGGAA CATCTAGTCG TTCAACAAAA TTGGTTCACG | 8220 |
| GAGGACTTCG TTACCTCAAA CAATTGACG TAGAAGTGGT CTCAGATACG GTTCTCGAAC | 8280 |
| GTGCAGTGGT TCAACAAATC GCTCCACACA TTCCAAAATC AGATCCAATG CTCTTACCAG | 8340 |
| TTTACGATGA AGATGGAGCA ACCTTTAGCC TCTTCCGTCT TAAAGTAGCC ATGGACTTGT | 8400 |
| ACGACCTCTT GGCAGGTGTT AGCAACACAC CAGCTGCGAA CAAGGTTTTG AGCAAGGATC | 8460 |
| AAGTCTTGGA ACGCCAGCCA AACTTGAAGA AGGAAGGCTT GGTAGGAGGT GGAGTGTATC | 8520 |
| TTGACTTCGG TAACAACGAT GCGCGTCTCG TGATTGAAAA CATCAAACGT GCCAACCAAG | 8580 |
| ACGGTGCCCT CATTGCCAAC CACGTGAAGG CAGAAGGCTT CCTCTTTGAC GAAAGTGGCA | 8640 |
| AGATTACAGG TGTGTAGCT CGTGATCTCT TGACAGACCA AGTGTTTGAA ATCAAGGCCC | 8700 |
| GTCTGGTTAT TAATACAACA GGTCTTGGA GTGATAAAGT ACGTAATTTG TCTAATAAGG | 8760 |
| GAACGCAATT CTCACAAATG CGCCCAACTA AGGGAGTTCA CTTGGTAGTA GATTCAAGCA | 8820 |
| AAATCAAGGT TTCACAGCCA GTTTACTTCG ACACAGGTTT GGGTGACGGT CGTATGGTCT | 8880 |
| TTGTCTCCC ACGTGAAAAC AAGACTTACT TTGGTACAAC TGATACAGAC TACACAGGTG | 8940 |
| ATTTGGAGCA TCCAAAAGTA ACTCAAGAAG ATGTAGATTA TCTACTTGGC ATGTGCAACA | 9000 |
| ACCGCTTCCC AGAATCCAAC ATCACCATTG ATGATATCGA AAGCAGCTGG GCAGGTCTTC | 9060 |
| GTCCATTGAT TGCAGGGAAC AGTGCCTCTG ACTATAATGG TGGAAATAAC GGTACCATCA | 9120 |
| GTGATGAAAG CTTTGACAAC TTGATTGCGA CTGTTGAATC TTATCTCTCC AAAGAAAAAA | 9180 |
| CACGTGAAGA TGTGAGTCT GCTGTCAGCA AGCTTGAAAG TAGCACATCT GAGAAACATT | 9240 |
| TGGATCCATC TGCAGTTTCT CGTGGGTCTA GCTTGGACCG TGATGACAAT GGTCTCTTGA | 9300 |
| CTCTTGCTGG TGGTAAAAATC ACAGACTACC GTAAGATGGC TGAAGGAGCT ATGGAGCGCG | 9360 |
| TGGTTGACAT CCTCAAAGCA GAATTTGACC GTAGCTTTAA ATTGATCAAT TCTAAAACTT | 9420 |
| ACCCTGTTTC AGGTGGAGAA TTGAACCCAG CAAATGTGGA TTCAGAAATC GAAGCCTTTG | 9480 |
| CGCAACTTGG AGTATCACGT GGTGTTGGATA GCAAGGAAGC TCACTATCTG GCAAATCTTT | 9540 |
| ACGGTTCAAA TGCACCGAAA GTCTTTGCAC TTGCTCACAG CTTGGAACAA GCGCCAGGAC | 9600 |

| | | | | | |
|------------|------------|------------|------------|------------|-------------------|
| 914 | | | | | |
| TCAGCTTGGC | AGATACTTTG | TCCCTTCACT | ATGCAATGCG | CAATGAGTTG | ACTCTTAGCC 9660 |
| CAGTTGACTT | CCTTCTTCGT | CGTACCAATC | ACATGCTCTT | TATGCGTGAT | AGCTTGGATA 9720 |
| GTATCGTTGA | GCCAATTTTG | GATGAAATGG | GACGATTCTA | TGACTGGACA | GAAGAAGAAA 9780 |
| AAGCAACTTA | CCGTGCTGAT | GTCGAAGCAG | CTCTCGCTAA | CAACGATTTA | GCAGAATTAA 9840 |
| AAAATTAAGA | AAAAATAAAA | GAGGTGGAGG | GCAGCATTCC | TTGTGCGCCG | TCCCTTCTTT 9900 |
| TTAATGGAGA | CAGAAAGATG | ATGAATGAAT | TATTTGGAGA | ATTCTAGGG | ACTTTAATCC 9960 |
| TGATTCTTCT | AGGAAATGGT | GTTGTTGCAG | GTGTGGTTCT | TCCTAAAACC | AAGAGCAATA 10020 |
| GCTCAGGTTG | GATTGTGATT | ACTATGGGTT | GGGGGATTGC | AGTTGCGGTT | GCAGTCTTTG 10080 |
| TATCTGGCAA | GCTCAGTCCA | GCTTATTTAA | ACCCAGCTGT | GACCATCGGT | GTGGCCTTAA 10140 |
| AAGGTGGTTT | GCCTTGGGCT | TCCGTTTTGC | CTTATATCTT | AGCCCAGTTC | GCAGGGGCCA 10200 |
| TGCTGGGTCA | GATTTTGGTT | TGGTTGCAAT | TCAAACCTCA | CTATGAGGCA | GAAGAAAATG 10260 |
| CAGGCAATAT | CCTGGCAACC | TTCAGTACTG | GACCAGCCAT | CAAGGATACT | GTATCAAACCT 10320 |
| TGATTAGCGA | AATCCTTGGA | ACTTTTGTTT | TGGTGTGAC | AATCTTTGCT | TTGGGTCTTT 10380 |
| ACGACTTTCA | GGCAGGTATC | GGAACCTTTG | CAGTGGGAAC | TTTGATTGTC | GGTATCGGTC 10440 |
| TATCACTAGG | TGGGACAACA | GTTTATGCCT | TGAACCCAGC | TCGTGACCTT | GGACCTCGTA 10500 |
| TCATGCACAG | CATCTTGCCA | ATTCCAAACA | AGGGAGACGG | AGACTGGTCT | TACGCTTGGA 10560 |
| TTCTCTGTGT | AGGCCCTGTT | ATCGGAGCAG | CCTTGGCAGT | GCTTGTATTC | TCACTTTTCT 10620 |
| AGTTTATACT | CTTCGAAAAT | CAAATTCAAA | CCACGTCAGC | GTCGCCTTAC | CGTACTCAAG 10680 |
| TACAGCTTGC | GGCTAGCTTC | CTAGTTTGCT | CTTTGATTTT | CATTGAGTAT | TAGAAAACAA 10740 |
| TTATGTTGAT | AGAGCTTGGG | CAAGAGCCCA | ATTTCAGCAA | AAAATGAAGT | AAATCTTCTC 10800 |
| ATAATAAAAC | GCATCATATC | AAGCACGAAA | ATCCACGAG | GTCAACTACA | GTCAGAAAGC 10860 |
| TGAACAACAA | GCCAAAACGC | CCAAAAAAGG | CGGCAAAAAG | CAAGCACCTG | CAAGCAACGT 10920 |
| GCCGAAATGG | TCAAATCCTG | ATTATGTCAA | CGAATTAGAC | CCAAAAATCG | TTGATATGCT 10980 |
| AGTAGAATTT | CACAAGTCAC | AAGGCACTTT | GGAAACTCCC | GAGGCGCAAG | CAGAAATCGC 11040 |
| CCAAAAACGT | GAAGAAATCG | AGCAAAGGAG | AGCTGAGCTT | GAGGGTAAAA | AACAAGAGCT 11100 |
| TTTGAACCGC | TTGAACAAAT | AGAGTTTCGC | AAGTATTATG | CTTACAAATT | ACTTGAGCAA 11160 |
| TTAACTAAAA | TATAAACCTT | GCCTTTATAT | CTAGGCAGGG | TTTATATTTT | AGAAATTCAC 11220 |
| GTAGGTTGTT | ACGGTTTTTA | CATACCCAGT | ATAGTTTGAG | TTTCTATAGT | ATTCACTGAT 11280 |
| AAACTTCCAT | TTTCTTTGAG | CAACATGGAT | ATAAGTACTT | GTTATGTAGT | ATGGATATGG 11340 |
| GCTTTGTGAA | TCCAAGTAAG | ACTGATAAGC | TTGTATACCA | AAATATGCTC | CACCAATTAT 11400 |

915

| | |
|--|-------|
| TGCACCCCAT GGACCCCCCA ATAAAGCACC TATCCTACCA ATCATATAAC TGATTCCAGC | 11460 |
| ACCAGTCATG AAGTTAGCGA ATGTGTTAGC TTGTTTATTC CCATGTATTG TGTGACGTA | 11520 |
| ATTCCAAACA TTAGGATCGT ATGATCTAAA AGATATATTT AGGTCGATT CATTCCTTTG | 11580 |
| ATAAGCCATA TAAAATGCCC CATTGATATA GACGCCGTCA GCACGTCGTT CAATAGTGTC | 11640 |
| TACTACTTCCA TCTGGATTGA CAACCTCAAG AACTTCATCG CTTAAAATAT TTAATTGCGT | 11700 |
| ATCTCCGAAC CGCACTGATG AGCCATTCTC AAACGAGCC TCACCAGATA CAACTTTAGA | 11760 |
| GTTTGCCGAT AAGCTATCAT CAGCAAAAAC AAACAAGCGA CGGGGAAATG CTAGACATAC | 11820 |
| AGAAAACAGA CATAACTAGC AAACACATGC ATTTAAACAT CTTAGACATA ACGGAAACTC | 11880 |
| CTTTGTATTT TTGATTTTTT TCAACTTTTA TTATACAATA AAACCAAATA AAAAGAAAGC | 11940 |
| GGTAACAATA TGCTTAATGC GAAAATTTTT TATATATTTT TATGTTTGAT CGTTATCGAA | 12000 |
| ACTACAGGCT TGTGTGTGTT GAAAAGAGGT CTCGAAATGG GTTATTTAGA CACAGAAGCT | 12060 |
| ATTATCCTCG CAGTTTTTTT ATTTGCTTTT TACAACCTAT GTTCATTTCGC TTGGGTCTGC | 12120 |
| TCTACAATAA AAAACAATAA AAAATAAATA GACGTATTTT CAAAAAAAC maAATGCATA | 12180 |
| TTTATATTAG CAAAACGACG ATTTAAATCG TCGTTTTTTT GTAGTACGAC GGGCATGTCC | 12240 |
| TATATCTGAG GTGTAAGTCC TCAGCCTGAC TATCGTGAGG TAGCAGGGAG AGGAAGGGAT | 12300 |
| AGCGAAATCG TGGCTCTACG AACAGGAACG TGATAGTAAG GCGTATATAG CCGATAAGGA | 12360 |
| GGCTTCAAAC TCTAAAGTCC AAAAAGGTAG TCGTAACCTA TATGTGTAAA TCACGAGAGT | 12420 |
| AATTGAATTC GGAATAAGGT TTGTGTGAAA AAGATAAATC TTTCTAGAGT CTAAGAGACTC | 12480 |
| TGCGTCAGAT TTCCTATTTT CACTGTAACC TTTTAACGTC CTCATATCTT GTATAAACGA | 12540 |
| GGAAAGATGT ACGACTTATC CCGTGAGGTT TCATGAGCGT GAAAGCGTAG TAACAACGAA | 12600 |
| TCATGAGAAG TCAGCCGAGC CCATAGTAGT GAGGAACTT CCGTAATGGA AGTGGAGCGA | 12660 |
| AGGGG | 12665 |

(2) INFORMATION FOR SEQ ID NO: 135:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 5305 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 135:

| | |
|---|----|
| CGCTAATCAC TACAATCATT TTATTGTACT TTTTCACTCT CAAGAAAAGC AAGAAGTATT | 60 |
|---|----|

916

| | |
|--|------|
| CATTTTAGTT TCATTTAGTA TTATTTTGCA TACCTAAAAAT ACAGTAAAAA ATCAGTCATC | 120 |
| TTGGTATGCT CCTGCTTTCA CTATTCAACA CGTTTTTGAC TTATACTAGG CTCATTTCCA | 180 |
| AAAGCATTAT ATAATAGTGA TATGAAACCA ACTAACTAA ACAAGAAATA TAAGCAATAA | 240 |
| AAATTCGTTT AAAAGATCTT ACTAAAGCTA ATACTAAATA AAAATAAAAG AGTAACTAG | 300 |
| GAAGTTTATT TCAAACAACC TAAAATACTG ATTTTCGGCT GAAGATAATA CTGGAGTGCA | 360 |
| AATTAATGGG GTTATAATAA ATAGCTGATA GCTTGTGTTG GTTTTGGATT TTTTAAGAGT | 420 |
| AGATGAGTAT TAAACTATA AGGAGGACGA AGGTGGCTAA AAATTTAAAA TTAAATTTAG | 480 |
| CTCGGGTAGA GCGTGATTTA ACACAAGGTC AACTGGCAGA GGCTGTCGGG GTGACACGCC | 540 |
| AGACTATTGG TTTAATAGAG GCGGGAAAAT ACAATCCCAG TCTCTCGCTC TGCCAGTCTA | 600 |
| TTTGACAGATG TTTAGGGAAA ACCCTAGACC AACTATTTTG GGAGGAAGAA GATGAAAAAT | 660 |
| AGATTTTATT ATTCTCAATT ACTAGACGAA AGAGAAGAAC AACTGTTCAA TAAAGCGGGC | 720 |
| TCTGAAAGTT TCTATATCTG CATTGCTTTG TCGCTCCTAT CTTATATCAT TTCAGTATTA | 780 |
| GCACCAAGCC TTTTAAATTC TAATATGCTG CTAATCGTTA TCATCATAGG GACATTTTAC | 840 |
| TTTTTCAATC GTGCCCGTTA TCTGGGAGTG ACCTACTATG GTCGTTTTCA TTTTACGATT | 900 |
| TTGGGTGTT TTTTCCTAAC CTGGCTATT ACGGCTCTTT TGATGTTGCA GAATTATCAA | 960 |
| TTCAACATAG AAATTTATCA GCACAATCCT TTGAATTTTA AATACCTGTC TGCTTGGGTC | 1020 |
| ATTACTTATA TCATTTACCT TCCGTGGATC TTTATTGGCA ATCTTGGTCT TAAGAGCTAT | 1080 |
| GGCGAATGGG CTCAGAAAAA ATTTGAACAA GATATGGATG AATTGGAGAG TGGAGAATAG | 1140 |
| CTTGTTACTC TTTTCTCAAT CCAGCTAAAA TGTGATATAA TAGTACTAAT TTATTGGAAT | 1200 |
| ACATGAAAGT TCTTGAAAAA TTTCATGGGT TTCTAGCTAA GGAAGTAGGA AAAGTATGTA | 1260 |
| TCCAGATGAT AGTTTGACAT TGCACACGGA CTTGTACCAG ATCAACATGA TGCAGGTTTA | 1320 |
| CTTTGACCAA GGGATTCAACA ATAAGAAGGC GGTCTTTGAG GTGTATTTCC GCCAACAGCC | 1380 |
| TTTTAAGAAC GGCTATGCGG TTTTGCAGG TTTAGAAAGA ATTGTGAACT ATCTTGAAGA | 1440 |
| CTTGCGTTTT TCAGATAGTG ATATAGCCTA TTTGGAGTCG CTTGGTTATC ATGGGGCGTT | 1500 |
| CTTGGATTAC CTTGCAATT TCAAGTTGGA GTTGACCGTT CGTTCTGCCC AAGAAGGGGA | 1560 |
| TTTGGTTTTT GCTAATGAAC CGATTGTGCA GGTGGAAGGA CCTCTAGCCC AATGTCAGTT | 1620 |
| GGTCGAAACG GCTCTTTTGA ACATCGTCAA CTACCAGACT TTGGTGGCGA CGAAGGCAGC | 1680 |
| TCGTATTCGT TCGGTTATCG AAGATGAACC CTTGATGGAG TTTGGGACAC GTCGGGCTCA | 1740 |
| AGAAATGGAT GCGGCCATCT GGGGAACACG CGCAGCTGTG ATTGGTGGCG CCAATGGAAC | 1800 |
| CAGCAACGTG CGTGCGGGTA AGCTCTTTGA CATTCCTGTT TTGGGAACCC ATGCCCATGC | 1860 |

917

| | | | | | | |
|------------|------------|------------|-------------|-------------|------------|------|
| CTTGGTACAG | GTTTATGGCA | ATGACTATGA | AGCTTTCAAG | GCTTACGCTG | CGACCCACAA | 1920 |
| AAATTGTGTC | TTTCTTGTGG | ATACCTATGA | CACCCCTTCGC | ATCGGTGTAC | CAGCTGCCAT | 1980 |
| TCAGGTGGCG | CGTGAGCTGG | GTGATCAGAT | TAACTTTATG | GGTGTGCGGA | TTGACTCTGG | 2040 |
| GGATATTGCC | TACATTTCTA | AGAAAGTCCG | TCAGCAACTG | GATGAGGCTG | GATTTACAGA | 2100 |
| GGCTAAGATT | TATGCTTCTA | ATGATCTAGA | TGAAAATACC | ATCCTTAACC | TCAAGATGCA | 2160 |
| AAAGGCCAAG | ATTGATGTCT | GGGTGTGGG | TACCAAGCTG | ATTACAGCCT | ATGACCAGCC | 2220 |
| GGCTCTTGGG | GCGGTTTACA | AGATTGTTGC | AATCGAAGAT | GAAACTGGTC | AGATGCGCAA | 2280 |
| TACGATTAAG | CTGTCTAATA | ATGCTGAAAA | AGTTTCTACG | CCAGGTAAGA | AGCAGGTGTG | 2340 |
| GCGCATTACC | AGTCGTGAAA | AAGGCAAGTC | AGAAGGCGAC | TATATCACTT | ATGATGGTGT | 2400 |
| GGATATTAGC | GACATGACAG | AAATCAAGAT | GTTCCATCCG | ACCTATACAT | ACATCAAGAA | 2460 |
| GACGGTTCGT | AATTTTGATG | CCGTTCTCT | CTTGGTGGAT | ATCTTCAAAG | AAGGAATATT | 2520 |
| AGTTTACAAC | TTGCCTAGTT | TGACTGACAT | TCAGGATTAT | GCCCCGTAAGG | AATTGACAA | 2580 |
| GTTGTGGGAT | GAGTATAAGC | GTGTGCTCAA | TCCGCAGCAC | TATCCAGTGG | ATTTGGCGCG | 2640 |
| TGATGTATGG | CAAGATAAGA | TGGACTTGAT | TGATAAGATG | CGCAAGGAAG | CCCTTGGTGA | 2700 |
| AGGAGAAGAA | GAATGAGTTT | GCAAGAAACG | ATTATCCAAG | AGCTGGGTGT | CAAACCAAGT | 2760 |
| ATTGATGCCC | AGGAAGAAAT | CCGTCGTTCT | ATTGATTTCT | TAAAAAGATA | TCTGAAAAAA | 2820 |
| CATCCCTTCC | TAAAAACCTT | TGACTAGGG | ATTTCTGGGG | GACAAGACTC | AACCTTGGCA | 2880 |
| GGACGTTTGG | CGCAATTAGC | TATGGAAGAA | CTGCGAGCTG | AAACGGGAGA | CGATAGCTAC | 2940 |
| AAATTTATCG | CTGTCCGCCT | GCCATACGGA | GTGCAAGCTG | ATGAAGCAGA | TGCTCAAAAA | 3000 |
| GCCCTAGCCT | TCATCCAGCC | AGATGTCAGC | TTGGTTGTGA | ATATCAAGGA | ATCAGCTGAT | 3060 |
| GCCATGACAG | CTGCAGTTGA | AGCGACAGGT | AGTCCTGTTT | CAGACTTCAA | CAAGGGGAAT | 3120 |
| ATCAAGGCAC | GTTGCCGTAT | GATTGCTCAG | TATGCCCTTG | CTGGTTCCCA | TAGCGGAGCG | 3180 |
| GTCATTGGAA | CAGACCACGC | CGCGGAAAAT | ATCACAGGTT | TCTTTACCAA | GTTTGGTGAC | 3240 |
| GGCGGTGCGG | ATATTCTCCC | TCTTTACCGC | CTCAATAAAC | GCCAAGGAAA | ACAGCTCTTG | 3300 |
| CAGAAACTTG | GCGCAGAGCC | AGCCCTTTAT | GAAAAAATCC | CAACGGCAGA | CCTAGAAGAA | 3360 |
| GATAAACCAG | GCCTAGCTGA | CGAAGTCGCA | CTTGGAGTCA | CCTACGCAGA | GATGACGAC | 3420 |
| TACCTAGAAG | GCAAAACAAT | CAGCCCAGAA | GCTCAAGCGA | CCATTGAAAA | CTGGTGGCAC | 3480 |
| AAAGGCCAAC | ACAAACGCCA | CTTACCCATC | ACCGTATTTG | ATGACTTTTG | GGAGTAAAAA | 3540 |
| GGTCCGGGGG | ACCTTTTATG | CTTCTTGCCC | TGAAATTAAA | AAGCAAGAAA | AACCTCCACT | 3600 |

918

| | | | | | | |
|-------------|-------------|------------|-------------|-------------|-------------|------|
| GGAGGTTTTC | AGCCTCTCAT | CTTGAAATAA | GAAAGTGAGA | GAAGGTCTGG | GGGATCTTGA | 3660 |
| ACCCCAGATT | TAGAAATAAG | AAAATGAGGC | AGATTCAGTA | ACTCGAAGAG | TTCGATTTCA | 3720 |
| TCGTCTTACC | CCTGCAACGA | TGACTAGGTT | TGAAAAAGCT | TGCTAGAGCG | CATTTCAAAC | 3780 |
| CAGGCAGCAA | CTGCGTCAAG | AAATTAGAAG | ACAAACTCGT | TTTCTAGCTG | TTACTGAGTT | 3840 |
| GAGCCTTTTT | ACTACGAGTA | TAGAAATAAG | GAAGTGAGGT | AGCATCATGA | AATCTATCGG | 3900 |
| TACGCAAATA | TTACAGACAG | AACGTTTGAT | TTTAAGAAGA | TTTGTGGAGA | GTGATGCAGA | 3960 |
| AGCCATGTTT | CAAAATTGGG | CTTCATCCGC | TGAGAATCTG | ACCTATGTTA | CCTGGGATCC | 4020 |
| CCATCCTGAT | GTCGAAATCA | CTCGAACTC | GATTTGCAAT | TGGGTTGCTT | CCTATACTAA | 4080 |
| TCTCAACTAT | TATAAATGGG | CCATTTGTCT | AAAAGAAAAC | CCAGAGCAAG | TAATAGGAGA | 4140 |
| TATCAGCATT | GTTAAGATAG | ACGAGGCTGA | TTTAAGCTGT | GAAATTGGCT | ATGTGTTAGG | 4200 |
| CAAGGCTTAC | TGGGGAAATG | GTATGATGAC | AGAGACTTTG | AAAGCTATCT | TGGACTTTTG | 4260 |
| TTTTACTCAA | GCAGGTTTTC | AAAAGGTCAG | AGCAGGTTAT | GCCAGTCTCA | ACCCAGCTTC | 4320 |
| AGGTCGTGTC | ATGGAAAAGG | CTGGAATGTC | CTATCTACAA | ACCATTGTTA | ATGGTGTAGA | 4380 |
| GAGAAAAGGC | TATCTTGC GG | ATCTTATTTA | TTATGGTATA | AGTAGGGAAG | AATGTTGAAT | 4440 |
| TCTATTTTCT | GTTTCTATCG | AAGTCAACTA | TTTATTGTAA | ATATAATAAT | TAGCATTC CA | 4500 |
| AGTTTATTTG | AAACTTTAAA | ATAGCATATT | GATTAGTACA | AGACAGATGT | TCTAGTTCCT | 4560 |
| TCTTTAATCT | GGTTTAGTGT | TAGTTAAAAA | ATCGCTTTAA | GCTTGTA ACT | AAGAGGGAGC | 4620 |
| TAATCGACTA | GATTCTCCAG | CCGAACAGGI | GGTAATGTAC | TTTTTATAGT | GTAATCCTAG | 4680 |
| CTGTTGT TAA | ATTTAAAATA | GAATCCTCTA | TCGAGTTAGG | GAATTAAATT | CAACCAATTT | 4740 |
| TATTCATGTT | TTTTCTATCA | AATTATCTAA | TATTAAAAATA | GTCTCATTCT | GATGAGAAAA | 4800 |
| CTATTC CCAA | ATCATTCATA | CCTCTCTCAA | CTAGATGTAA | CTTACAAAAC | CCCTGACCTC | 4860 |
| ATGAGCCACT | TTCTTCCTCC | TCATGAGGTC | AGTTTTACTT | TCTGCTGTTC | CAGTATCGTT | 4920 |
| TTTCCTCGCT | AGATTTCCTC | AAAAGGGCAG | ACTCCTCCCT | TGGTGCGTCA | CACGATTTT T | 4980 |
| TCATCTCGAC | TGTTCTTTAA | TGCATCATTA | ACGACGCTTT | TCTTCTAGGT | GGTTCATAAG | 5040 |
| GAACAGGAAG | ATTCAGGTTG | ACTTTTCTAA | TCCTAGAATA | AAGTGCTGAA | AACAATTCGG | 5100 |
| AATAGGCATA | GAGACTAGAC | AATTTGAGGA | GCTGCTTGCG | TCCTGTTCGA | ACACATTTTC | 5160 |
| CCACCACGTG | AAGAAAAAGA | TGGCGGAAGC | GTTTGATTGT | TAAAGTTTGG | AAGTCACCTC | 5220 |
| CAGCTAGATG | TTTGAGAAAA | AGATAGAGAT | TGTAGGCGAT | ACAGCTCATC | ATCATACGAA | 5280 |
| CTTCGTTTTT | GATTAAGGTT | GAACT | | | | 5305 |

(2) INFORMATION FOR SEQ ID NO: 136:

919

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3964 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 136:

| | |
|--|------|
| TGGCAGCTCG TCGTCGTAAA GGACGCAAAG TTTTGGCTGC ATAATCCAAA CGAATTCTAT | 60 |
| CAAAAATCAG TAGGAACTCG AGTCTACTGA TTTTATTTT TGTA AAAAAG TTCAGTAGAT | 120 |
| GCAAAATGGAT TCGGAAGCGA TGTTACAGTA GATTGAAACT AGAATAGTAC ACCTCTGTTT | 180 |
| CTAA AACATT GTTAGAAATC GATTGACTG TCCTGATCGA TTTGTCCTGT TATTATTTTA | 240 |
| TTT TACTATA AAGTTGAAGT AGGTGGAGAT GGTACAGCAA CAATCGTCTT TAAAGATGGT | 300 |
| TCAGCTATTA CAATTCCAGG AAATCAATTG GTAGCACAAG ATCCAAAAGC ACAAGATAGC | 360 |
| ACTAAACTGA CTGCTGAAAA ATCAACTGTT AAAGCACCTG CTCAAAGAGT AGATGTAAAA | 420 |
| GATATAACTC ATTTAACAGA TGAAGAAAA GTTAAGGTTG CTATTTTACA AGCAAATGGT | 480 |
| TCAGCATTAG ACGGAGCGAC AATCAATGTA GCTGGAGATG GTACAGCAAC AATCACATTC | 540 |
| CCAGATGGTT CAGTAGTGAC GATTCTAGGA AAAGATACAG TTCAACAATC TGCGAAAGGT | 600 |
| GAATCTGTAA CTCAAGAAGC TACACCAGAG TATAAGCTAG AAAATACACC AGGTGGAGAT | 660 |
| AAGGGAGGCA ATACTGGAAG CTCAGATGCT AATGCGAATG AAGGCGGTGG TAGCCAGGCG | 720 |
| GGTGGATCAG CTCACACAGG TTCACAAAC TCAGCTCAAT CACAAGCTTC TAAGCAATTA | 780 |
| GCTACTGAAA AAGAATCAGC TAAAAATGCC ATTGAAAAAG CAGCCAAGGA CAAGCAGGAT | 840 |
| GAAATCAAAG GCGCACCCTT TTCTGATAAA GAAAAAGCAG AACTTTTAGC AAGAGTGGAA | 900 |
| GCAGAAAAAC AAGCAGCTCT CAAAGAGATT GAAATGCGA AAATATGGA AGATGTGAAG | 960 |
| GAAGCAGAAA CGATTGGAGT GCAAGCCATT GCCATGGTTA CAGTTCCTAA GAGACCAGTG | 1020 |
| GCTCCTAATG CTGCTCCTAA GACAACAAGT GCACCGCAAG CAACTGCAGG AACAAATGCAA | 1080 |
| GATGTTACCT ACCAGTCACC TGCTGGCAAA CAATTACCTA ACACAGGTTC AGCATCAAGT | 1140 |
| GCAGCACTTG CTAGTCTTGG TCTAGTGGTG GCAACAAGTG GTTTTGCTTT GCTAGGAAGA | 1200 |
| AAGACTAGAC GTAGAAAATA GAACAGCTAG AAAATTCTAT TCTCTACTTA AAGTTAGATT | 1260 |
| ATAAGGGGGA TTTTGAGAAG TCATCAATCC TAGTGATGGG TGAGAAAAAGT GAGAACCCAA | 1320 |
| GATAATCACA TACTTTAGCT GAATAGGAAT ATTCTATCAA TGAGCCAAT CTCTTCTGTC | 1380 |
| TCTAACTGTG GAATAGGAGA TGGGCAATAT CGGATAGAAA AGATAGCAGA ATAGCTCTCT | 1440 |

920

| | |
|--|------|
| ATTGAAGAGA GGAGGGGAAA CCGAAAAATT AGGTGCCCCCT CCTCTTTTTT GGTATAATAG | 1500 |
| AAGATAGAAA ACGAGGTTAG AAGAGATGAT TTTTGATACA CATAACACT TGAATGTAGA | 1560 |
| AGAATTGCA GGTGCGTAGG CAGAAGAAAT TGCCTTGCT GCTGAGATGG GTGTGACACA | 1620 |
| GATGAATATT GTTGGTTTG ATAAACCGAC GATTGAGCAT GCCTTGAGT TGGTAGATGA | 1680 |
| GTATGAGCAG CTCTATGCGA CTATTGGTTG GCATCCTACA GAAGCTGGTA CTTATACAGA | 1740 |
| GGAAGTTGAG GCTTACTTGT TGGATAAGTT AAAACATTCC AAGGTTGTGG CTTTAGGTGA | 1800 |
| AATTGGCTTA GATTACCATT GGATGACAGC GCCCAAAGAG GTGCAGGAGC AGGTTTTTCG | 1860 |
| CCGTCAGATT CAGCTATCTA AGGACTTGGA TTTGCCTTTT GTTGTCCTATA CCCGTGATGC | 1920 |
| GCTGGAAGAT ACCTATGAGA TTATCAAGAG TGAGGGCGTT GGTCTCGTG GTGGTATCAT | 1980 |
| GCATTCATTT TCAGGGACGC TTGAGTGGGC AGAGAAGTTT GTGGATCTTG GTATGACCAT | 2040 |
| TTCTTCTCA GGAGTGGTGA CTTTAAAGAA GGCAACTGAC CTCCAAGAAG CAGCTAAAGA | 2100 |
| GTTACCTTTG GACAAGATGT TGGTGAAAC AGATGCGCCT TACTTAGCAC CTGTACCCAA | 2160 |
| GCGTGGTCGT GAAAAATAAA CAGCCTATAC TCGCTATGTG GTCGACTTTA TCGCTGACTT | 2220 |
| GCGTGGTATG ACGACAGAAG AGCTGGCGGT AGCAACGACT GCAAATGCAG AACGAATTTT | 2280 |
| TGGACTGGAC AGCAAGTAAT GAAAGAGAAA ATTTCTCAAG TTATCGTGGT TGAAGGCGT | 2340 |
| GATGATACGG TCAATCTCAA ACGTTATTTC GATGTGGAGA CCTATGAGAC TCGAGGTCT | 2400 |
| GCCATCAATG CTCAGGATAT AGAGCGGATT CAGCGCCTGC ACCAACGTCA TGGAGTCATT | 2460 |
| GTCTTTACAG ACCCAGATTT TAATGGGGAA CGGATTCGGC GCATGATCAT GATGGTCATT | 2520 |
| CCAACAGTTC AGCATGCCTT TCTCAAGCGA GATGAAGCTG TTCCCAAGTC CAAGACCAAG | 2580 |
| GGGCGTTCTC TGGGAATTGA GCATGCCAGC TATGAAGACC TGAAAACGGC TCTAGCTCAA | 2640 |
| GTGACAGAAC AATTTGAACA TGAGAGTCAG TTTGACATTA GTCGTAGCGA TTTGATTCGC | 2700 |
| CTTGGTTTTT TAGCAGGGGC AGACAGCCGT AAGCGTAGAG AATATCTCGG AGAGACTCTC | 2760 |
| CGAATCGGCT ATTCCAACGG CAAGCAACTC CTCAAACGCC TAGAGTTGTT TGGGGTACT | 2820 |
| TTGGCAGAAG TGAAGAAGC TATGAAATCT TATGAGTAGG AAAGATGTAG CCGTTACAAT | 2880 |
| TTTTTAAGTT TCACAGTATT TTTCGAAGCA GGTAAGAG GAGGCGTCTG ATGTTAATTG | 2940 |
| GTCAAAAAAT TAAAGAGATT CGGATAGAAA AAGGAATTAG TCGTCCAGAT TTTTGTGGAG | 3000 |
| ATGAGCAAGA ACTGACAGTT CGTCAACTGT CGCGAATTGA AAGTGGAGCT TCGCAACCGA | 3060 |
| GTTTGCCCAA GTTAGACTAT ATTGCTCGCC GGCTAGGAGT TCCAGTTTAT AGCCTTATGC | 3120 |
| CGGATTTTTC AGCTCTTCCT TCTGCTTATT TAGAATTGAA ATACCAGATT TTACGTGAAC | 3180 |
| CAATCTATGG TAAAGAAGAG GAGTACGATA AGAAGGAAGC GTGTTTGAA GAGATTTATA | 3240 |

921

| | |
|--|------|
| AAACATACTT TGATAATCTT CCTAAAGAAG AACCAATTAGC ATGTGAAGTA TTGCAGGCGT | 3300 |
| GTTTGGATAC TTCTAGAACT AGAAGGCCTG AATATGCAGA GTTAATACTT GAGGAACATA | 3360 |
| TGCCTCAGAT TATAGAAAAA GAAGCTTATT CAATAAATGA TATGTTGTTG ATTCGTTTGT | 3420 |
| TTTTTTATCA AATGCTCATT AGAAAAGATC TTGCCAAATT TATAAATCAA ATCGAAAAGC | 3480 |
| TAATGCTCTT TCTTTTGGAA CAGAAGAAGG TAACTCAAAT AGAGAATTAC TTTATAATTA | 3540 |
| GAGATACTCT TATTTCAGGA ATGTGTTGTC TTGAAAAGGT AGGAGTAACT GATTGTTTTA | 3600 |
| ATGATTATCT ATCGTGTTTA CAAGAAATTA TGGATAAAAC TCAAGATTAT CAAAAGAAAC | 3660 |
| CTCTGTGATT TATGTTTTTG TGAAGCAAG CATTAAGAGA AGAAAGAGAT TTTAGTTTAG | 3720 |
| CTGAATCATT TTATCAGTCT TCTAAACAT TTGCGCAGCT AATTGGAGAT GAATTTCTAG | 3780 |
| TAAAGAAATT GACAGAGGAA TGGCAAGAGG ATGTCAAAAA ATATTTATAA ACATAGTGAA | 3840 |
| TCAGTGACAA AGATGTCCTT GTCCTCGTAT CAAAACAGTT CTAAAGTTCG TCTTTAGGGA | 3900 |
| TGTTTTTTTA GATATAAGCT AAAAATGACA CGAAATGGTT AGATTTTAAG GACATTGATG | 3960 |
| TCCG | 3964 |

(2) INFORMATION FOR SEQ ID NO: 137:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 12666 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 137:

| | |
|--|-----|
| TGAGACCGTT ATTTGTATTA GGGAAATGGG TATCTATTTT TAATGCTGTG GGGATTTTGA | 60 |
| TTGTTTCTAT TATTCAAACC AAAAGCTTGT CAGGTATTGG AGCAGGATTG TTAAATCTAT | 120 |
| ATAACATTTC ATCTTATATA GGTGATTTAG TTAGTTTCAC TCGATTGATG GCATTAGGAT | 180 |
| TATCTGGAGC AAGTATAGCA TCAGCTTTCA ATTTAATTGT TGGTTTGTTC CCGGGAATAT | 240 |
| TGGCTAAACT GACAATTGGA TTAGTATTAT TCATTCTTTT ACATGCCGATC AATATTTTTC | 300 |
| TATCGTTACT ATCAGGATAT GTTCATGGAG CACGTCTGAT ATTTGTTGAA TTTTGTGGTA | 360 |
| AGTTTATGA GGGTGGAGGA AAACCATTTT AACCTTTGAA GGCTTCTGAG AAATATATTA | 420 |
| AGGTATTAC AAAGAATTAA TGGAGGATAT ATATAATGGA ACATTTAGCA ACTTATTTT | 480 |
| CAACCTATGG AGGAGCTTTC TTCGCTGCAT TGGGAATTGT ATTGGCGGTT GGATTAAGCG | 540 |
| GTATGGGGTC TGCTTATGGA GTTGGTAAGG CTGGGCAATC TGCCGCAGCT TTAATGAAAG | 600 |

922

| | | | | | | |
|------------|-------------|------------|-------------|-------------|-------------|------|
| AACAGCCTGA | AAAGTTTGCC | TCAGCTTTGA | TATTGCAATT | ATTGCCCGGA | ACACAAGGAT | 660 |
| TATATGGTTT | TGTTATTGGA | ATTTTAATTT | GGTTGCAATT | AACTCCAGAA | CTTCCTTTAG | 720 |
| AAAAAGGCGT | TGCTTATTTT | TTTGTAGCTC | TTCCAATTGC | TATTGTAGGA | TACTTTTCAG | 780 |
| CTAAGCATCA | AGGAAATGTA | GCAGTAGCGG | GAATGCAAAAT | CTTGGCTAAA | AGACCAAAAAG | 840 |
| AATTCATGAA | GGGAGCAATT | TTAGCTGCCA | TGGTAGAAAC | CTATGCAATT | CTTGCTTTTG | 900 |
| TCGTATCATT | CATTTTGACC | CTTCGTGTAT | AAGAAATAAA | TTTGCAATTC | AAAGGAGGTG | 960 |
| TCTAAATGAG | CAATTTAGAA | AACTTACGAG | AGTCTGTTAT | TGAACAAGCT | CATGAAAAAG | 1020 |
| GGCGTATGAA | ATTATTGGAT | TCCAAAAAGA | AGATTGATGA | TGAATTTGAA | ATGCAAAAGT | 1080 |
| CGCTCATTAT | AAAGAAAAAA | GAAGCTGAAC | ATGAACGAAA | GTAAAAAGAA | TTGCAACAGA | 1140 |
| AATATCAAAT | AATTTTTCAA | CAATTAAAAA | ATAAGGAACG | CCAATCAACG | TTAGTATCAA | 1200 |
| AACAGAAAAT | ATTAAAAAGAA | CTTTTTCAT | CTGCTTTACT | AGAAATGGAA | TCTTGGAGTG | 1260 |
| CAGATAAAGA | AATGGAGTTC | ATCTATCGAA | TTCTGGAACG | ATATTCACAA | CAAGAGGTCA | 1320 |
| TAGTAACCTT | TGGGGAACGG | ACTTTAGCTA | AATTCAAATTT | GGAACAATTA | GAGAAATTGA | 1380 |
| AATTCCTCTT | TCCAAATTAT | TTATTTAGTG | AACAACCTAT | CTCAAAATGAA | TCAGGCTTAC | 1440 |
| TTATTTCAAT | AGGTAAAAAT | GATGATAACT | ATTTGTATAA | AACATTAATT | GGATCGATTT | 1500 |
| CTAAGGAAGA | AAGTTCAAGT | ATCGCAAATC | AAATTTTAT | CAATTAAGGA | TGAAATGGT | 1560 |
| TAATCCTTCT | TAGAAATTTG | GAGTATTCCA | ATAAAATTAG | AAAGGTATTT | TATGGATACT | 1620 |
| AATCTTTTTT | CAAAAAATAAA | TACGACGATT | TCGGTAAAAG | AAAACGATTT | TATTACAGAA | 1680 |
| GAAAAATTTT | AAAAAATTAT | ACAATCCAAA | GATACGGAGA | CATTGGCATT | TATCTTAGAA | 1740 |
| TCAACTCCCT | ATCATTTATC | GATTGACATC | TTAGAAGATC | CTAGTCAGAC | AGAGATTTCT | 1800 |
| CTAATGACAA | AATTAGTCAA | TGATTATAGA | TGGGCCTATG | CTGAAAGTCC | GTCTGATATA | 1860 |
| ATTGTGACTT | TATTTGCTTT | ACGATATGTT | TATCATAATA | TCAAAGTTTT | ATTAAAATCT | 1920 |
| AAGGCGGCAA | TTAAGAAAGA | TTTTTCTAAA | TTATTAATTC | CAATAGGGAT | TTTTGATATA | 1980 |
| GAAAGTTTAA | AACATTTAGT | TTCTTCCTTA | CATTCAGATA | CACTTCCTGA | TTTTATGGTT | 2040 |
| CGTGAAGTAG | AATCAATTTG | GAATGAGTAT | GAAACTTTTA | ATAATATTCG | TGTACTTGAT | 2100 |
| GTCCGAGCTG | ATCTAGCATA | TTTTAAACAT | CTGAAACTTT | TATCTAATGA | GTTAGATGAG | 2160 |
| GTACTGTCTC | AGGTATTGTT | CGAAATGATT | GACTTTTATA | ATATTATTAC | TGTAAAACGT | 2220 |
| GGTTTATCTC | AAAATAAGAG | TCATGGGGAT | ATTTTACAAT | TACTTTTCAGA | TGAAGGAAGT | 2280 |
| ATTTCTGCTA | AAGAAATTTAT | ATACATTGTA | GAAAAATCAAG | AAATATTGTT | GTGGTTCAAT | 2340 |
| AAAATAAATC | CAAGCTTAGA | TTCAATCTTT | TCAACTTATG | AATTGAAGAT | GCAGGACGCA | 2400 |

923

| | |
|--|------|
| ACAATTTTCAT CTTCTGAGTT AGAATTTTFA TGTGATTTAC TATTGTATAA AACTTTAGAT | 2460 |
| CAAGGAAGGT ACAATGTAGA GGGGCCGTTA GTTCTTGCTA GATATTTATT GGGATGTGAG | 2520 |
| TTTGAAGTAA AGAATCTCAG AATGATCATA TCAGCTCTTC AAAATACAAT TCCCTTTGAA | 2580 |
| TCAATAAAAG AAAGGATACG CCCACATTAT GGAAGCTAAT AAGTATAAAA TTGGCATAAT | 2640 |
| TGGTAGCCGT GATATTATTT TACCATTAG CATGATTGGG TTTGATATAT TTCCTGCCTA | 2700 |
| CCAAGAACAA GAAGCTATAA ATACACTAAG AAAATTAGCT CAATCTGATT ATGGTGTCAT | 2760 |
| TTATATCACT GAAGACATTG CTTCAATGAT ATTAGATACA ATTCGCCATT ATGATTCCCA | 2820 |
| AGTTGTGCCT GCTATTATTT TATTACCGAC TCATAAACAA GGTTTAAAT TAGGATTAAA | 2880 |
| ACGTATAGAG GATAATGTAG AGAAAGCAGT AGGACACAAT ATTTTATAAT AATGTACAAA | 2940 |
| ATTGTCGTGA ATATTATCT ATAATTTTG GACTTAGTAA GGAGAATAAC TTTGACTCAA | 3000 |
| GGGAAGATTA TAAAAGTATC GGGACCTCTA GTTATTGCAT CAGGTATGCA GGAGGCTAAT | 3060 |
| ATTCAAGATA TTTGCCGTGT AGGTAAGCTA GGGTTAATCG GTGAAATTAT TGAAATGAGA | 3120 |
| AGAGATCAGG CATCTATCCA AGTCTATGAA GAAACATCTG GTCTTGGTCC GGGAGAACCT | 3180 |
| GTTGTTACAA CTGGAGAACC TCTCTCGGTT GAATTAGGGC CAGGATTGAT TTCTCAAATG | 3240 |
| TTTGATGGCA TACAACGCCC ATTAGATCGA TTTAAATTGG CTAATCATAA TGATTTTCTA | 3300 |
| GTTCGTGGGG TAGAAGTTCC AAGTTTGGAT AGAGATATTA AGTGGCATTT TGATTCCACT | 3360 |
| ATAGCAATTG GTCAAAAAGT GAGTACGGGT GATATTCTTG GAACTGTCAA GGAAACCGAG | 3420 |
| GTAGTTAATC ATAAAATTAT GGTTCCTTAT GGAGTATCTG GAGAAGTCGT TTCTATTGCA | 3480 |
| TCTGGCGATT TTACAATTGA TGAAGTTGTA TATGAAATAA AAAAATTGGA CGGTAGTTTC | 3540 |
| TATAAAGGAA CGCTTATGCA AAAATGGCCT GTCCGCAAGG CGCGTCCTGT TTCTAAACGT | 3600 |
| TTAATTCCAG AAGAACCATT AATCACAGGT CAACGAGTTA TTGATGCATT CTTTCCAGTA | 3660 |
| ACCAAAGGGG GAGCTGCAGC AGTTCCTGGA CCGTTTGGAG CAGGAAAGAC AGTTGTACAA | 3720 |
| CACCAAGTAG CTAAATTTGC CAATGTTGAT ATTGTTATTT ATGTCGGTTG TGGAGAACGT | 3780 |
| GGAAATGAAA TGACGGATGT ACTGAATGAG TTTCTGAGT TGATTGACCC TAATACCGGA | 3840 |
| CAATCAATTA TGCAACGGAC AGTTCTGATT GCTAATACTT CAAATATGCC TGTGCTGCT | 3900 |
| CGTGAGGCTT CAATTTATAC AGGAATTACC ATGGCTGAGT ATTTTCGTGA TATGGGCTAC | 3960 |
| TCTGTCGCCA TTATGGCTGA TTCAACTTCA CGTTGGGCAG AAGCGCTACG TGAAATGTCA | 4020 |
| GGACGTCTAG AAGAAATGCC TGGTGATGAG GGTATCCTG CTTATCTGGG AAGTCGTATC | 4080 |
| GCTGAATATT ATGAAAGAGC AGGACGTTCT CAGGTTCTAG GGCTTCCAGA ACGTGAAGGA | 4140 |

924

| | | | | | | |
|------------|------------|-------------|-------------|-------------|-------------|------|
| ACGATTACTG | CTATTGGAGC | TGTATCGCCA | CCTGGTGGAG | ATATTTTCAGA | ACCAGTFACT | 4200 |
| CAAAACACTT | TACGGATTGT | GAAAGTTTTT | TGGGGGCTTG | ATGCTCCGTT | GGCACAGCGA | 4260 |
| CGTCATTTTC | CTGCAATTAA | CTGGCTTACA | TCTTATTCAC | TATATAAAGA | CAGTGTGGGC | 4320 |
| ACTTATATAG | ATGGTAAAGA | GAAGACAGAT | TGGAATAGTA | AAATAACTCG | TGCGATGAAC | 4380 |
| TACTTACAAC | GGGAATCTAG | TTTAGAGGAA | ATTGTTTCGTC | TTGTTGGAAT | TGATTCTCTG | 4440 |
| TCTGATAATG | AACGACTAAC | GATGGAAATT | GCTAAACAAA | TTGAGAGAAG | TTATTTGCAA | 4500 |
| CAGAACGCTT | TTGATTCGGT | AGATACATTC | ACTTCGTTTG | CAAAACAAGA | AGCAATGCTA | 4560 |
| AGTAATATTC | TCACTTTTGC | TGATCAGGCA | AATCATGCTT | TAGAGTTGGG | TTCTTACTTT | 4620 |
| ACAGAGATTA | TGGAAGGTAC | CGTGGCAGTT | CGAGACCGTA | TGGCGAGAAG | TAAATATGTT | 4680 |
| TCAGAAGATA | GATTAGATGA | AATCAAAATT | ATATCAAAATG | AGATTACACA | TCAAATTCAT | 4740 |
| TTGATATTAG | AAACAGGAGG | TCTATAAATG | AGTGTATATA | AAGAATACAG | AACTGCTAGT | 4800 |
| GAAGTTGTTG | GGCCTCTTAT | GATTGTTGAA | CAAGTAAATA | ATGTGTCTTA | CAATGAGTTA | 4860 |
| GTTGAAATTC | AACCTCATAA | TGGAGAAATT | CGTCGTGGAC | AAGTTTTAGA | GATCCACGAA | 4920 |
| GATAAAGCAA | TGGTTCAGCT | TTTTGAAGGA | TCTAGTGGAA | TAAATTTAGA | AAAGTCTAAA | 4980 |
| ATTCGTTTTG | CTGGTCATGC | ATTAGAATTG | GCTGTATCTG | AGGATATGGT | TGGTCGTATT | 5040 |
| TTTAATGGGA | TGGGAAAACC | AATFGATGGT | GGACCAGATT | TAATTCCAGA | GAAATATTTA | 5100 |
| GATATTGATG | GTCAAGCTAT | TAATCCTGTA | TCTAGAGATT | ATCCAGATGA | ATTTATTTCAG | 5160 |
| ACAGGGATCT | CCTCTATTGA | TCATTTGAAT | ACTCTTGTA | GTGGTCAAAA | ATTACCAGTA | 5220 |
| TTTTCAGGTT | CGGGCTTACC | TCATAATGAA | TTAGCTGCTC | AGATAGCAAG | ACAAGCGACT | 5280 |
| GTTTTAAATT | CTGATGAAAA | TTTTCGGGTT | GTATTTGCAG | CAATGGGTAT | TACTTTTGAA | 5340 |
| GAAGCTGAGT | TTTTTATGGA | AGAACTCAGA | AAAACAGGAG | CGATCGATCG | TTCGGTTTTA | 5400 |
| TTTATGAACT | TGGCAAATGA | TCCTGCAATT | GAGCGTATTG | CAACTCCCCG | CATTGCTTTA | 5460 |
| ACTGCGGCAG | AGTATCTAGC | TTTTGAAAAA | GATATGCACG | TTCTAGTTAT | CATGACGGAT | 5520 |
| ATGACTAACT | ATTGTGAAGC | GTTACGTGAA | GTCTCGGCAG | CTCGCCGTGA | AGTTCCAGGG | 5580 |
| AGACGAGGCT | ATCCGGGATA | TTTATATACA | AATTTATCAA | CTCTATACGA | AAGGGCTGGT | 5640 |
| CGCTTAGTTG | GTAAAAAAGG | TTCCGGTGACA | CAGATTCCTA | TTTTAACAAT | GCCAGAAGAT | 5700 |
| GACATAACAC | ATCCAATTCC | TGATTTAACT | GGATACATTA | CTGAAGGGCA | AATTATTTTG | 5760 |
| TCGCATGAGT | TGTATAATCA | AGGTTATCGT | CCACCAATCA | ATGTTTTACC | TTCTCTCTCT | 5820 |
| CGATTAAAAG | ATAAGGGATC | TGGAGAAGGT | AAAACCTCGT | GAGATCATGC | TCCAACATATG | 5880 |
| AATCAACTGT | TTGCAGCCTA | TGCCCAAGGG | AAAAAGGTTG | AAGAGTTAGC | AGTAGTATTA | 5940 |

925

| | |
|--|------|
| GGAGAATCGG CTTTATCTGA TGTAGATAAA TTGTATGTGA GGTTTACAAA GCGTTTGTGAA | 6000 |
| GAAGAGTACA TAAACCAAGG ATTTTATAAA AATCGAAATA TAGAAGATAC GTTGAATCTT | 6060 |
| GGGTGGGAAT TACTATCAAT TCTTCCTAGA ACAGAGTTAA AACGTATCAA AGATGATTGT | 6120 |
| CTTGATAAAT ACTTACCTTT GGTAGAAGTT TAATCCGGAA ATGGAGTGAT TATCTATGGT | 6180 |
| ACGTTTGAAT GTAAAACCAA CTCGTATGGA ATTGAATAAC TTAAAGGAAC GTTTGACAAC | 6240 |
| AGCTGAACGT GGACATAAGT TATTAAAGGA TAAAGAGAT GAATTGATGA GGCGATTAT | 6300 |
| TTCTTTGATT CGTGAGAATA ATCAACTTCG GAAAGAAGTG GAAAGTTATC TAATTGATAA | 6360 |
| TCTAAATCC TTTGCAGTTG CTAAATCATT AAAGAATTCT CAAATGGTGG AGGAATTATT | 6420 |
| TTCAATTCCA TCGAAAGAAA TTGAATTATT TGTGAGAAA GAAAATATCA TGAGTGTAAC | 6480 |
| AGTTCCTAGA ATGCATATGA ATATTACTTC TCAAAATGAG AACAGTGAAT ACAGCTATTT | 6540 |
| ATCTTCTAAT AGTGAAATGG ATGATGTATT TGCTACAATG AATAGTTTAA TTTATAAATT | 6600 |
| ACTAAGACTG GCAGAAAGTTG AAAAAACGTG TCAGTTAATG GCTGATGAAA TAGAAAAAAC | 6660 |
| ACGTAGACGT GTAAATGGTT TAGAATACTC GATTATTCCA AACTTGTCGG AACTATATCA | 6720 |
| TTATATAGAA TTGAACTAG AGGAGGCAGA AAGAGCCAAT TTAGTTCGTA TTATGAAAGT | 6780 |
| GAAGTAGATC CTTTATTTAG ATTATTAATT AGATGAACAA ATATCAGCTT GGATAAGGCT | 6840 |
| TTAAGCCTTT CTAAGCTTTT TTTATTGACA GTATCAGGAT ATCTTTTCA AAATTTGGT | 6900 |
| TTGTTAGATA ATGAAAATGT TTCTACTAAT CTAGATTAG GATTAGTAAA TCGTAAATGT | 6960 |
| AATTATATAG AAAGTAAGCG CGTCATAACA AGGTATCTAT CATTATGGA GCTCCTCTG | 7020 |
| TATACTATTA GTAAAGTAAA ACTATTGGAG GATATTTTAA TGCCACAACC TATTGTTCTT | 7080 |
| GTAGAGATTC CACAATCTCG TCGTTTGTAT TCTAAAAAGA GAAATGATAT TCTGCTTAAA | 7140 |
| ATTCGTATTG GCAAGCTTGA AGTAAGTTT TTTCAATCTC TCAATCTCGA AATGGTAGAA | 7200 |
| CAGCTTTTGG ATAAGGTGTT GCTCTATGAC AATTCATCTA TCTAGCCTAG GGGAGGTCTA | 7260 |
| TCTCGTGTGT GGGAAAACGT ATATGAGACA AGGAATCGAT TCACTGGCTT ATCTGGTTAA | 7320 |
| AACCCACTTT GAATTGGATC CTTTCTCCGG TCAAGTCTTT CTCTTTTGTG GTGGACGTAA | 7380 |
| AGACCGCTTT AAAGTCCTTT ACTGGGATGG TCAAGGATTT TGGCTACTAT ATAAACGCTT | 7440 |
| TGAGAACGGC AGATTGATTT GGCTAAGTAC AGAAAAGGAT GTCAAAGCTC TCACACCAGA | 7500 |
| ACAAGTAGAC TGGCTTATGA AGGGCTTTTC TATCACTCCA AAAATATAGT AGATTGAAAC | 7560 |
| TAGAATAGTA CACCTCTGCT TCTAAAACAT TGTTAGAAAT CGATTTTACT GTCTGATCG | 7620 |
| ATTTGTCTG TTCTTATTTT ATTTTACTAT AAATCCATCA GAAAGTCGTG ATTTCTATTG | 7680 |

926

| | |
|--|------|
| AAATGAGGAC TTTCTTTTTA TACTCATCTG CTTTCAAAAA GCATTCTAGT CCATCTCCGA | 7740 |
| TTAACGATGG ACTTTATCAC CTCCTTCTCC AGTCCTTGTA TAACATCTTG GAGTTGATTC | 7800 |
| ATGACATCTT CCAAAGTTTA AAAGGCTTTA TTCTTAAATC CACGTTTACG AATCTCTTTC | 7860 |
| CACACTTGTT CAATGGGGTT CATCTCTGGT GTGTATGGAG GAATAAATGC AAAGCCAATA | 7920 |
| TTAGTCGGAA TCTTTAAGGT ACTTGATTTA TGCCATATAG CATGTGCCAT AACGAGTAA | 7980 |
| AGATAATCAT CTGGATAAGC TTGTGAAATC TCCTATTCCT AAAGCCCCCT TAGCGCATAA | 8040 |
| CTTTGGCTCA GCTTCTATTA TCGCTCACAC CATCCATCAG AAGTTTAAATC TGAAGGTACC | 8100 |
| CAATTATCGC CAAGAAGAAG ATTGGGCTAG GATGGGTTTA CCAATCACAC GTAAGGAAAT | 8160 |
| CTCTAATTGG CATATCAAGG CGAGTCAATA CTATTTGGAG CCCCTTTATA ACCTCTTGCG | 8220 |
| AGAGAGACTA TTGACTCAGC CCTTACTTCA TGCGGATGAA ACTTCTTATA GGGTGCTAGA | 8280 |
| GAGTGATAGT CAGCTGACTT ACTATTGGAC TTTTGTGCA GGTAAGCAG AGAAACAAGG | 8340 |
| GATTACGCTT TACCACCATG ATCAGTGTGCG AAGTGGTTCA GTAGTACAAG AATTCCTAGG | 8400 |
| AGATTATTCT GGCTATGTGC ATTGTGATAT TTTGCGGCAG TAACTTAGGA CTTTAGTCCT | 8460 |
| CTAGTTCTGC CTATGCGATA GCAGTCCAAG GTTTAGGAGC AAGGCGACGC TAAGCTTGGT | 8520 |
| AAACTTCGAA CCGCTCGTCT GCTTATCGTC AACTGGAAGA AGCTGAACTT GTTGGATGTT | 8580 |
| GGGCGCATGT GAGAAGGAAG TTTTGTGAAG CGCCCCCA AGCAAGCGGA TAAATCATCC | 8640 |
| TTAGGAGCTA AAGGTTTAGC TTATTGTGAT CAGTTATTTT CCTTGGAAAG AGACTGGGAG | 8700 |
| GCTTTGCCAG CTGATGAACG ACTACAGAAA CGTCAAGAAC ATCTCCAGCC CTTAATGGAA | 8760 |
| GACTTCTTTG CTTAGTGCCG GCGTCAGTCA GTTTTAGCAG GTTCAAACT AGGAAGGCA | 8820 |
| ATTGAATACA GCCTCAAGTA TGAAGAAACC TTAAAGACCA TTTTGAAAGA CGGACATCTG | 8880 |
| GTCCTTTCCA ATAATCTAGC TGAACGCGCC ATTAAATCAT TGGTTATGGG ACGGAGTAAA | 8940 |
| AGAGTCCAGT GGACTCTTTT AGCCTAAGCT CAGTTTAAAA AAGCGAGGGT GGTATTTTTC | 9000 |
| TCAAAGTTTT GAAGGAGCTA AAGCAAGAGC TATTATTATG AGTTTGTGG AAACAGCTAA | 9060 |
| ACGTCATCAA TTAAATAGCG AGAAATATCT ATCCTATCTT CTAGAATGTC TTCCAAACGA | 9120 |
| GGAAACTCTC GTAAACAAAG AGGTTTGA GAAGTTATTTA CCATGGACTA AAGTTGTACA | 9180 |
| AGAAAAGTGC AAATAAGAAA TCTCCAGATT AGGAACTATC CGTGAGTTCT CCAGTCTGGA | 9240 |
| GATTTTTCAT TAGACTTCCT GCGAAACAAA ATATGGTATA ATAGTTCTAT GAATGATGAA | 9300 |
| GCAAGTAAAC AACTAACCGA TGCACGATT AAGCGTCTTG TTGGTGTCA ACGCAGGACT | 9360 |
| TTTGAAGAGA TGTTAGCTGT ATTAAAAACA GCTTATCAAC TTAACACGC AAAAGGTGGA | 9420 |
| CGAAAACCTA AATTAAGTCT AGAAGACCTT CTTATGGCCA CTCTTCAATA TGTGCGAGAA | 9480 |

927

| | |
|--|-------|
| TATCGAACTT ATGAACAAAT TCGGGCTGTT TTTGGTATTC ACGAAAGCAA CTTAATCCGT | 9540 |
| CGGAGCCAAT GGGTTGAAGT AACTCTTGTT CAAAGTGGTG TTACGATTTC AAGAACTCCT | 9600 |
| CTCAGTTCTG AGGACACGGT AATGATTGAT GCGACGGAAG TAAAAATCAA TCGCCCTAAA | 9660 |
| AAAAGAATTA GCGAATTATT CTGGTAAAAA GAAATTTTAC GCTATGAAGG CTCAAGCGAT | 9720 |
| TGTCACAAGT CAAGGGAGAA TTGTTTCTTT GGATATCACT GTGAACATTT GTCATGATAT | 9780 |
| GAAGTTGTTT AAAATGAGTC GCAGAAATAT CAGACAAGCT GGTAAAATCT TGGCTGACAG | 9840 |
| TGGTTATCAA GGGCTCATGA AGATATATCC TCAAGCACAA ACTTCACGTA AATCCAGCAA | 9900 |
| ACTCAAACCG CTAACAATTG AAGATAAAGT CTATAACCAT GCGCTATCTA AGGAGAGAAG | 9960 |
| CAAGGTTGAG AACATCTTTG CCAAAGTAAA AACGTTTAAA ATGATTTCAA CAACCTATCG | 10020 |
| AAATCATCTA AACGCTTCGG ATTACGAATG AATTTGATTG CTGGTATTAT CAATCATGAA | 10080 |
| CTAGGATTCT AGTTTTCAG GAAGTCTATT ATCAAAAAATA CCATCAAGAT TATATAAGAT | 10140 |
| TGATACAGGA AAAGTTTAT TTGATGGTGT AAATATTAAT CAAATAGATA AAAAAATATT | 10200 |
| AAGTCAAAAT TTAGGAGTAG TTCCACAGGA TTCATTTTTA TTGAACCGAA GTATTCTTGA | 10260 |
| TAATATAACT TTAAAGCACG AAGTTACTTC ACAAAGATA GAGGAAGTTT GTAAAGCAGT | 10320 |
| TCAAATCTAT GATGAAATCA TGGCTATGCC GATGAAATTT AATACTATCA TCTCAGAGAT | 10380 |
| GGGGTCAAAT ATTTAGGTG GGCAGAGCA ACGGATAGCA CTGGCACGTG CATTAATAAA | 10440 |
| TAATCCTAGT ATTGTAATTT TAGATGAAGC AACTAGTGCA TTAGACACTA TTAATGAGGA | 10500 |
| AAGAATAACA AAGTATATAC AAAGTCAGGG CTGTACTCAA ATAATTGTAG CTCATAGATT | 10560 |
| GTCAACGATT AAGGATGCGG ATGTTATTTT TGTAATGAAA GGTGGTAAGA TTGTTGAGTC | 10620 |
| AGGAAATCAT AAGTACTTAA TGGATCTTGG TGGAGAGTAC TACAGCTTAT ATACAAAAAG | 10680 |
| GAAATGAGGT GTAAAGAAAA TGAAGAAAGA AAATGAATAT GTAATTTTAA CAACAGCCTC | 10740 |
| ACTAGGGGTG ATGATTGGAA TAGTGTTCG AATTTTTTTA GATTTTCCAG TTGAATATGG | 10800 |
| TATTTCTTTA GGCTTGTTGA ATGGAATAGT ATTGGGTTTCG CTGATTGTTT AAAAAACAA | 10860 |
| TAAGAATTAA GCATAATTTT TTGCTGTAAA CTAAGGAGTA GAGATGGCTA TAGTTGAAAT | 10920 |
| TATAAATCTA AAAAAAGCT TTAAAGATAT TGAAGTTATT CATAACACTT AAATAATAGA | 10980 |
| GCAACTACAG TAGTAGCTTA AAAACATGAT TAAATCGCTA TTCTTAGGAG TAGCGGTTTT | 11040 |
| TCTTTTGTGTT TAATACTCTT TGAAAACTC TTCAAACCAAC GTCAGCTTTG CTTTACCGTA | 11100 |
| CTCAAGTACA GCCTGCGGCT CGCTTCCTAG TTTGCTCTTT GATTTTCATT GAGTATAAAA | 11160 |
| AGGGTCAAGT AAGTATAGTA AATTGAAATA AGATATGAAC AAATCGATTA GAAAAGTCAA | 11220 |

928

| | |
|---|-------|
| ATTAATTTCT AGAAATATGT TAGAAATTGG TTTGAATTCC GCAATCAATT TGTTCAGTTT | 11280 |
| TTATTTTCATT TCATTTTATT TAATTAGATT TTCCAATTTT TTAATTCAAG CTA AAAATCC | 11340 |
| CCAATCGTAG TGATTGAGGA TTGAGTAAAT AAATCTTAAA CAATACCTTG TGCAATCATG | 11400 |
| GCATTTGCTA CATTTTCAAA GGCAGCAATG TTAGCTCCTG CAAGGTAGTC TTTATCAAGA | 11460 |
| CCGTATGTTT CTGAAGTCGT TTTAGCTGTG TTGAAGATGT TTGTCATGAT GTCTTTGAGA | 11520 |
| CGGCCATCAA CTTCTTCACG AGTCCATGAG AGGCGAAGAC TGTTTTGGCT CATTTCAAGA | 11580 |
| GCTGAAACGG CTACACCACC AGCGTTGGCA GCTTTTGACG GTCCGTAGAA GATACCATTT | 11640 |
| TCTTTGTAAA CTTTGATGGC ATCAAGGTCG CTCGGCATGT TGGCACCTTC AGATACACAG | 11700 |
| ATAACGCCCTT GAGCAACCAA ACGTTTAGCT GCTTCACCGT TGATTTTCGTT TTGAGTGGCA | 11760 |
| CATGGAAGAG CAATGTCATA GTTCCACGCG TAAGTCCATA CAGTACCTTC GTGGTAGGTT | 11820 |
| GCAGTTGCTT TTTCAGCTGC ATACTCAGTC AAACGAGCAC GACGTTTTTC TTAAACATCA | 11880 |
| ACCAAAAGAT CGAAGTCGAT ACCATTTTCA TCGATGACAT AACCATTTGA GTCAGAAACA | 11940 |
| GAAATAACAG TTGCACCGAG TTCAGTTGCT TTTTGAAGAG CATATTGAGC AACGTTACCA | 12000 |
| GAACCTGAAA TAACGACTTT CTTACCAGCA AAGCTGTTAC CGTTAGCTTT GAGCATTTCT | 12060 |
| TCAGTATAGT AAACCAAACC GTAACCAGTT GCTTCTGGAC GAATCAAGCT ACCACCAAAT | 12120 |
| CCAAGAGGTT TACCAGTCAA GACACCAGCA TCAAATTGGT TAAGACGTTT GTATTGACCG | 12180 |
| TAAAGGTAAC CAATTTTCAG TCCACCAACA CCGATATCAC CAGCAGGTAC GTCAAGTGAT | 12240 |
| GGTCCGATGT GTTTTTGCAA TTCAGTCATG AAGCTTTGGC AGAAGCGCAT CACTTCAGCA | 12300 |
| TCTGTTTTAC CTTTAGGATC GAAGTCTGAT CCACCTTTAC CTCCACCGAT AGGAAGTCCA | 12360 |
| GTCAAGACAT TTTTAAAGAT TTGTTCAAAT CCGAGGAATT TCAAGATCCC TTGGTTTACA | 12420 |
| GTTGGGTGGA AACGAAGTCC ACCTTTGTAT GGTCCAACAG CTGAGTTGAA TTGAACACGG | 12480 |
| TAACCACGGT TTACTTGAAT TTTTCCATCA CGGTCAACCC AAGGAACACG GAAAGAAACC | 12540 |
| ACGCGCTCAG GCTCAGTAAT ACGTGCCAAG ATATTTTCTT CGATATACTC AGGGTGTTTT | 12600 |
| TCAAATACAG GTTCTAAAGT GTTGAAAAAT TCTTCAACAG CTTGGAGGAA TTCAGCCTCG | 12660 |
| TGCCCG | 12666 |

(2) INFORMATION FOR SEQ ID NO: 138:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3083 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

929

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 138:

| | |
|--|------|
| AGCAACTGTT GTGAACCAAT TCCGATAAAT TCCAAGAATT GGTAAATAGA GCCATTTTGA | 60 |
| CCAAAAATCC CGATAAAAGC ATAGGCTTTA AGGAGCAAAT TGATCCAGGT AGGAAGGATA | 120 |
| ATCAGCATGA GCCAGAGTTG ACGGTGTTTG AGACGGGTCA AAAAGAGGGC CGTCGGATAA | 180 |
| CTGATAAGCA GTGCCACAAA GGTCACAATG CCTGCATAAA GCACTGAGTT GAAACTCATT | 240 |
| TTAAGATAGG TCAAGTTTGG TGACGCAAAG TAAGATTTGT AATTTTCTAA ACTGAACTGG | 300 |
| CCTTCGATGT TGAAAAAGGA TTGACCGAAA ATCAAGACCA AGGGTGCCAA TACAAAGAGC | 360 |
| GCAATCCAAA GCATGTAGGG TACTACAAAG AGTTTAGAGC TTGTTTCTCT CATCTCTTTC | 420 |
| CTCCTCGATT GCATTGATCA AACCTGCTTC TTGCTCTTCG ATTTCTACGT ACTCCTCAAT | 480 |
| ACGAGCATCG AACTCTTCTT CGGTTTCATT GAGACGCATG ATGTGGATGT CTTCTGGTTC | 540 |
| AAAGTCCAGA CCGATTTCCCT CACCCACGAT AGCCTTACGG GTTGAGTGGA TCATCCATTC | 600 |
| ATTTCCAAGT TCGTCATAGG CGATAATTTT ATAATGAACT CCACGGAAAA GCTGGGTATC | 660 |
| GACCTTAACT TGGAGCTTGC CTTCTTCAGG AAGGGTAATG CGCAAGTCCT CTGGACGAAT | 720 |
| AACGACCTCA ACAGGTTTCAT TTGGCTTCAT CCCACCATCA ACCGCTTCAA AGCGTTTGCC | 780 |
| GTTAAATTCG ACCAAGTAGT CCTCAATCAT GGTACCTGGC AAGATGTTTG ACTCCCCGAT | 840 |
| AAAGGTGGCA ACAAAGTGGT TGATTGGCTC ATCGTAGATG TCCACAGGGG TTCCAGACTG | 900 |
| GACAACTCTG CCATCATTCA TAACGAAAAT CCAGTCACTC ATGGCAAGAG CTTCTTCTCG | 960 |
| ATCGTGAGTG ACAAAGACAA AGGTAATGCC CAATCGTTGT TGTAATTCAC GCAATTCGTA | 1020 |
| CTGCATGTCT GTTCTCAATT TCAAGTCCAG CGCTGATAAA GGCTCGTCCA ACAAGACCAC | 1080 |
| ACGGGGTTGG TTGATGATAG CACGGGCGAT GGCCACACGC TGACGTTGTC CTCCAGAAAG | 1140 |
| TTTGCGGATG GAACGTTTTT CATAACCTTC CAACTGAACC ATCTTGAGAA CTTCCGCTAC | 1200 |
| ACGCTGCTCG ATTTCTTTCT TATCAATTTT ACGCAAGCGA AGTGGAAGG CAACATTTTC | 1260 |
| AAACACATTC ATATGTGGGA ACAAGGCATA GGATTGGAAG ACGGTATGTA CGTCGCGCTT | 1320 |
| GTGTTGTTGA ATATCATTGA TACGAACACC GTCTAGCATG ATATCTCCTG TCGTCGCATC | 1380 |
| CAGTAAACCT GCAATAATGT TTAGGATAGT TGATTTCCCC GAACCAGATG CACCTAGAAG | 1440 |
| GGTGTAGAAT TTCCCTTCTT CCAACTCAAA GTTGATGTCT TTGAGAACCT TGGTGTGCT | 1500 |
| GTCTTCAAAA ACTTTAGAGA CGTTTTTGAA TTCGATAATT GGCTTTTCA ATTGGCATAA | 1560 |
| ATTCTTCTT TTTCATAGAT TAACCGATCG GGGCTCTGTC AGGTCCCCAC TACCTCTTGC | 1620 |
| AGGGAGTAAA ACCACCTGCA TACATCTTCG CTACCGATAG GCTTTCACCC AAGATCCGGA | 1680 |

930

| | |
|--|------|
| CTTCTCTTTC AAGCGTAATA CCTGAGTGTT CCTTGACTTT TTCGATAACC GATTGGATCA | 1740 |
| AGTCCTCGTA GTCTTTGGCC GTTCCATCTG CGACATTGAT CATAAATCCT GCATGCTTTT | 1800 |
| CTGACACTTC TACGCCACCG ATACGATAGC CTTTCAAGCC AGCTTCTGAA ATTAACGTAC | 1860 |
| CTGCAAAATG CCCGACTGGA CGCTTAAAGA CCGAGCCACA AGATGGGTAT TCCAAAGGTT | 1920 |
| GCTTGAGTTC ACGTAGGTGC GTCAAGCGGT CCATTTCCTG CTTGATAACC TGATGGGTTC | 1980 |
| CTGGAGCTAG GGCAAATTTA ACTGACAAGA CAACTGCACC AGACTCCTGA ATAGCTGAAT | 2040 |
| GACGGTAACC AAAAGCCAAG TCTTTAGCAG ACAGGGTTTC GATTTCCTCA TCCTTGGTCA | 2100 |
| AGACCTTACA AGACTGCAAG ATGTGAGCAA TCTCGCCACC ATAGGCACCC GCATTCTATA | 2160 |
| AGACAGCACC GCCAACGCTT CCTGGAATAC CACAAGCAAA CTCAAAGCCA GTTAAACTAT | 2220 |
| GACGGAGGGC AATGCGAGTT GTTTCAATCA AGTTAGCCCC AGCTTCTGCT TCAATGGTAT | 2280 |
| AGCCATCAAC AGAAACGTTA TTGAGCTTGT CACACAAGAT GACAAATCCA CGAATCCCAC | 2340 |
| CATCACGAAC GATGATATTG CTTGCATTGC CAAGAACCAT CCAAGGGATA TTTTCTTGGT | 2400 |
| TGGCAAATTT CACAACGCGA GCCAACTCAA AACGATTTTCG TGGAAAGACC AAATAATCAG | 2460 |
| CCTCTCCACC TACTTTTGTA TAACTATAGC TATGCAAGGG TTCCTTAAAA CGGATATCAA | 2520 |
| TTCTTCTTAA GATTTCAAGC ATTTTCTCTC TTACAGACAT GTCACCTCTC CTTTACAAA | 2580 |
| ATTCAATCCA TTATACCATT TTTAGAGACA TTTGACGACC ATAAAAATAC CTGTGTTTGA | 2640 |
| TTTTCATATA GAAAAAGAGG TTCCCCCTT TTTATGATTT TTTACAAAAG ATTTCTTGG | 2700 |
| TTCCATAGGC GACCAGAACG AGCTCCAGTG CTAGAATCAC TTCAACCAAG ACTGGATTG | 2760 |
| TCAACCAGCC TACTTGGAAG AGAGATGGTG CCAGATCAAA GAAGGCATGC AAGCCATAGG | 2820 |
| CTGCTAGGAG ATAAATCCAT TTCTTCTGGC GAACAGCTTG GTAAACCCAA ACTGTCAAAA | 2880 |
| GTAATTGGAA ACCAAGCGCC AAGATTCGCT CAAAACCAAG CAAATAAATC TGCCAGACCG | 2940 |
| AAAGTGAAGT AATGGTTTTT AACATATTTT CAGACAGTAA TTGCATAACC TGTGGATTC | 3000 |
| GAGTTTGAAC TGCCGAAAGA ACAATGTAAA GATTGAGTAA ACTAGTAAGG CCTAGAAAAA | 3060 |
| TCAACTCCAA GCCACCATGC CCC | 3083 |

(2) INFORMATION FOR SEQ ID NO: 139:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15363 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 139:

931

| | |
|--|------|
| CCGGAGGATA TTGACCACCA CCAAAGCAG GGGGAAAATC GAAATCAACC AATAGTAGGC | 60 |
| TACTGCGACA CTGGTCAACT CACTATCTGA TGCTTGATAA TAATGCAAAA AAGCTTTTAA | 120 |
| TAAAGGTTTG TCTATCAGCT CTTTCCACCA CTTTTCATG TCATACTCCT TCACCTATAA | 180 |
| TCTTATACTC AATGAAAATC AAAGAGCAAA CTAGAAAGCT AGCCGCAAGC TGCTCAAAAC | 240 |
| ACTGTTTTGA GGTGTAGAT AAGACTGACG AAGTCGATCA CATACATACG GTAAGGCGAC | 300 |
| GCTGACGTGG TTTGAAGAGA TTTTCGAAGA GTATTAACCTA ATTTCTTCTT ACCAATTCCA | 360 |
| CCATATCATA CGGTAGGGTA TTGGCAGCTT CCTTCAAGGA ATAGTTCTCT AAGTTATTTA | 420 |
| CATTTTGTCG TAATTTCTTG GCATACTTAG TCGTAATCAA TCGTTTTTCT TCGTATTCGA | 480 |
| AAATCAACTT GCGCTCCAGA TAATAGCCTC TCAGCATTTT ATCGATATTG TTGGGTTTGA | 540 |
| CACGATTGAT AACCCGTTTCG ACAAAGGCAC CACTGCTGAT AATAGCTGTT TCTCGAAGAC | 600 |
| GAGACTCCTG CATAAACTA ATCAAAGAGC GTCTGTAGAC TCCCTTCAGG TTTTCCAAAC | 660 |
| TTTCAATAAT CATCTCTGTA TTGGCAAGAT AGAGCTCTGC AATTGGTCA TAATCAAGAG | 720 |
| CACGGAGACG GCTTTGCTCC TTGTTCTTCC AGCTACGGAA GGTCTTTCG AGAGTAAAAA | 780 |
| CTTCATGAAG GAGAAAACGT AAAATCCTCA AGGAAACAAG AAAATAATAG GTCAGTCTTG | 840 |
| AGGCAAGTTT ACGATTGATT CCTTGTTCTA TATTTTTCAG ATAACGTTGG TAAACTCGGT | 900 |
| AAGCACGATT GCTAATGTTT CCCTCTTCAT AGGCCTGTTT CAAACCATCA CTTTCAATAC | 960 |
| TAAGAATCAA GAGTTTCAA GCAGCCCAGT CTTCTTGATC ATCCTGGTTT TCTTGGCTTA | 1020 |
| AAATGAGATT TTCAATACGT CCATGATAAT TGTCATAGC CGCATAGAGG GGAAGTTTAT | 1080 |
| TTCTGGTGTC TTCCAACCTT TTTTCCAAC CTAGCGTTAC TTCATTCAA ATGGCGATAT | 1140 |
| GCATAAGATA ATCCTTGCTT TCTTCTCTT CATCAGAAAG ATGAGGCAAG ACCAAGAGAC | 1200 |
| CTGTTAAAAA GCTAACAAGC GTCACACCTG CAACAAGGAA AAGCAAAAGA GGATACTCCT | 1260 |
| GTTCTAGATT ACTTGGTATC AAGAGAATCG TAGCAATCGA CACCGTTCCC TTAACACCTG | 1320 |
| AAAAGGTCAA GAGAAACATG TCCTTCATAT ACTTATTTAG CTTTTCCTTG AGGCGTCGGG | 1380 |
| TTCTATAGGC ATAATAGCCA TAGATCATAA TAAAACGAAT GACAAAAAGG ACAAAGGTAA | 1440 |
| GGGCGATAAG AGATAGCAAT AAAAGTAGAG GATTATAGAT TGGATTGGTC AAGATAGGTT | 1500 |
| CTGCTATCAT TTCCAACCTC ATCCCTAAAA TCACAAAGAC AGAACCGTTG AGCATAAAGG | 1560 |
| TCACTGTATG CCAGACCGTC TCGGTCACCG TATCCACTTG GGCTTCGAGG AGCGTGATTT | 1620 |
| TCTTGAAGCG ACTTGCCTTT AAAATTCCAG CAACTACGAC GGCAATAATA CCTGAAACAT | 1680 |
| GAACCTCTTC TGCCAGAAAG AAGGTCACCTA GAGGCAAACT CAATTCTAAT AAAAGTTCAC | 1740 |

| | | | | |
|-------------|------------|------------|------------|------------|
| 932 | | | | |
| TGGCAATATC | CGTTGCGCGC | ACACTTAGCA | AGAAGGTATG | GAGGAAGCGG |
| CTGTTAAAA | TCCAATTAAA | AAACCGCCTA | GGATTGAAAA | GATGAGCGAA |
| GCCCCAGAGA | AAAAGCTCCA | GTGTCCAAG | CTGTCAAAGC | TACCTGAAAA |
| CAGAAGCATC | ATTCAAGAGT | CCTTCGCCCT | TAAGAATATT | GGACACGCGC |
| TAAAACGCTC | CGAAAGAGAG | GCAAAGGCCA | CCAAGTCCGT | AGGACCAAGG |
| CAGCCAAGCA | AGCTGCCAAG | GGAAGGCTGA | ACCAAAGAAG | ATGGGCCAAG |
| TCAGGGTCGA | GATAAAAATC | ACTGGAAATA | TGAGATAAAC | AATGATTTCG |
| AAATAGCCGT | AACATCTGCT | TCTTCAGCCT | CTCGGAAAAG | CAAGGGTCCG |
| CCAAAAACAA | CTCCGTATTA | AGGTGAAAGT | CAGTATTGGG | TAAAAAGAGA |
| TTCCCAAAAG | AATTTCACAC | AAAGGGAGAG | GCAAAAAGGG | CAGGAGCTTA |
| TTGAGACAAT | CAAAACCAGT | AAAAATAGGA | TGAGGTAAAT | CAGTAATTCC |
| TCCTTAATCT | TTTTTACAAC | AGGATTCAAA | TATCTCCTTC | TGCTCTTTGA |
| AATCTTGGA | CAGTCTTTGT | GCTCAATTTT | TCTCTGGCAC | CGTTCCATTT |
| TAATTTTTTC | TTGATTTTAA | GCATTTTTTT | GCTCATATGC | GCTTGGTCTA |
| CGCTCGTTCTG | TGGTGGGTG | ATTCAACAAA | ATTCTGGCGC | ATGGCATCCA |
| TAAGTATGT | TTATCCATGT | CTGTATCTCT | CTAATTTTTC | AATCATCACT |
| GGTTGTTGAC | TTGGTTTAAA | GTTCGGTAAA | TGGCAGCTGT | GTAATCTTGT |
| GGATCACAAA | ATCCAAGACA | GCATCTCTCT | CGAGATCGCC | TCCTTCATGA |
| TCATAATAGC | AATTCGTCCA | CCTTTGACAA | GTAAGCCACA | TAGCTTTTCT |
| TCGTTGTCTG | CGGTCGGGTG | ATGACAGACT | TATCAGCTGC | CGGCAAATAG |
| AAATCCCTGC | CTTAGCTTTT | ATCACAAACT | GGTCCAGTGT | CTCATGGCCT |
| ACTGGGCATT | TGTCAAGTCA | GCCTGATGCA | AACGCTCTTG | GGTCTTTTCC |
| TCTGAATATC | AAAGGCATAG | ACTTGCTTGG | CTAGCTTGGC | TAAAAAAGC |
| CATTTCCCAT | AGTCGCATCC | ACTACGACAT | CCTCTTTTGT | CACGACCTCA |
| CATGTGCCAT | CTCAAGTGGT | CTTTTCATTT | TCAAACCTCT | GTTTACAGC |
| TGAACACTTC | CACGACGTCG | CATCTCCATC | TCAATGCTGT | TGAGGACTTC |
| AGGCTCCACA | TAGGACCAAG | CAGCATATCC | CTAGGCGCAT | CTCCTGTAAT |
| ACGATATGTT | TGGGAATAAT | TTCCAGTTGG | TCACAGATGA | CCCTGACATA |
| CTCATCAATT | GTAAACGCCC | CTCATGGTAA | TCTCGTTGCA | TACGAGTATT |
| TGGAGCAAAT | GCAGTTTAAT | CCCTTGAATA | TCGTTATCCG | TGACACAACG |

933

| | |
|--|------|
| TCAACCATCA TCTCATGGGT TTCACCAGGC AAACCATTGA TCAAATGGGA AACAACTCTCA | 3600 |
| ATTTTGGGAT ACTTTCTCAA ACGCTTGACC GTTTCACCT ACAATTCATA AGAATGCGCA | 3660 |
| CGGTAAATCA GGTCAAGAGT TGCTTCATAA GTAGTTTGCA AGCCCAATTC AACCGTCACA | 3720 |
| TGCATGCACT CCGATAAATC AGCCAAATAT TCGATGGTFT CGTCTGGTAA ACAGTCTGGG | 3780 |
| CGCGTTCCAA TATTGATTCC TACCACACCT GGCTCATTGA TAGCCTGTTC ATAACGCTCT | 3840 |
| CGAATAACTT CCACCTTTTC ATGGGTGTTG GTAAAAATTTT GAAAAATAAC CAGATACTTC | 3900 |
| CGAACATCCG GCCACTTGCG GTGCATAAAG TCAATTTTCCT TATAAAATTG CTCACGGATA | 3960 |
| GGCGCATCCG GTGCCACAAT GGCATCTCCA GAACCAGAAA CCGTACAAAA AGTACAGCCC | 4020 |
| CCATGAGCCA CAGTCCCATC ACGATTGGGA CAATCAAATC CCGCATCAAT AGGGACTTTA | 4080 |
| AAAGTCTTTT CTCCAAAGAG TTTTCGATAA TAATCATTCA AGGTATTATA AGATTTCATG | 4140 |
| ACTTTCATTA TAACAAAAAT CACCCACAAT CTCAAAAGCC TGACTTTCCT ATAAATTCCT | 4200 |
| CTGTTTCTCG TTTCCATTAG CCTTTTMTTA TGATACAATA TGGGTATGAT TTTAATGAAA | 4260 |
| TTAGCATCTA TTTTATTATT GATACTGACC TTAGTCGTCT GCATTATCCT AACCAAATTC | 4320 |
| TTTAGATTAA AAAAATAAGG ACGAACTTTT GCGGATTGG CTTTTCAGT CTTGGTATTT | 4380 |
| GAGTATTACT TGATTACAGC TAAAACCTTT ACCCATAATT TCCTCCCTAG ACTGGGGCTA | 4440 |
| GCCCTCTCGA TCCTAGCCAT TATTCTCGTC TTTTCTTCC TTTTGAAAAA ACGCAGCTTT | 4500 |
| TACTACCCTA AATTTATCAA ATTCTTCTGG CGTGCAGGAT TCTTATTAAC CCTTATCATG | 4560 |
| TATATAGAAA TGATTGTTGA ATTGTTCTTA ATGAAATAGT CGAATCCCTA AGCATTTTCT | 4620 |
| AGGGATTTTT GCTTCTCTA CAAAATAGTA TAGACAATAA CACTATACAA TTTTATACAA | 4680 |
| AGAAAAGAGT CTGGGACAAT AGTCTCTTAT ATCCAAAAAG GCAACGGATT TGCCGTTGCT | 4740 |
| TTTTTGATG GTTACGATAG TCTTGGTAAA ATAGAATTGC CCAATAAACC ATTTAGAAAG | 4800 |
| GCTATCCCAT GCATATTCAC TATAACACAA ATCAAACAAC TTTACCACTA GAAATCAGTT | 4860 |
| CCTTCTTACC ACAAGATCAT CTCGTTTTTA CTATTGAAAA AGTGGTGAAT ACCTTGGAGG | 4920 |
| AACGTCACTT CTACACCTCC TATCATGCCT TTGATCGCCC GTCTTATCAC CCTAAAATGC | 4980 |
| TTGTATCTAC TCTTCTATTT GCCTATTCAC AAGGGATTTT CTCTGGTCGA AAAATTGAAA | 5040 |
| AATGGAAGAG TTAGTGACCT TAGATTGTTT GTTTATTGAC AGAACTAAGA TTGAAGCCAA | 5100 |
| TGCCAACAAG TATAGTTTTG TGTGGAAGAA AACGACAGAG AAATCTCTCG CCAAATTCAT | 5160 |
| AGAACAGATA CAGGTCTATT TTCAAGAAGA AATCACTCCC CTTCTGATTA AATATGCCAT | 5220 |
| TTTGATAAG AAACAAAAGA GAGGTATATA AGAGTCAGCT AAAAATCTAG CGAATTGGCA | 5280 |

934

| | | | | | | |
|------------|-------------|------------|------------|-------------|-------------|------|
| CTATAATGAC | AAGGAGGATA | GCTACACACA | TCCTGATGGC | TGGTATTATC | GTTTTCACCA | 5340 |
| TACCAAATAT | CAGAAAACAC | AGACAGACTT | TCAACAAGAA | ATCAAGGTTT | ACTACGCCGA | 5400 |
| CGAACCTGAA | TCAGCCCCTC | AAAAGGGACT | GTATATGAAC | GAACGCTATC | AAAAC TTGAA | 5460 |
| AGCTAAAGAA | TGTCAGGCGC | TTTTATCTCC | CCAAGGTAGA | CAGATTTTCG | CTCAACGCAA | 5520 |
| GATTGATGTG | GAACCTGTCT | TTGGGCAGAT | AAAGGCTTCT | TTGGGTTACA | AGAGATGTAA | 5580 |
| TCTGAGAGGG | AAGCGTCAAG | TGAGAATTGA | CATGGGATTG | GTACTTATGG | CCAATAACCT | 5640 |
| CCTAAAATAT | AGTAAAATGA | AATAAGAACA | GGACAAATCG | ATAAGGACAA | TCAAAATCGAT | 5700 |
| TTCTAACAAT | GTTT TAGAAG | TAAAAGTGTA | CTATTCTAGT | TTCAATCTAC | TATACAATAA | 5760 |
| GAGAATGACT | CAAAATTAAA | AAGCTAGAGT | TCCACAATTG | GAAATATCTA | GCTTTTTTGT | 5820 |
| GGTTGAGAAC | TATTTTGTCT | CAGGCTCTTT | ATCTTCTATT | TAGGACAAGA | GTTTTCCTTT | 5880 |
| GGTCTTTAAT | GATAAAGAAG | GTATCAAAAT | TTCTAGTCTT | CTTTTTTACC | TTTAGTAACT | 5940 |
| ACTAATCCTG | CACTCAAACC | TAGAAGAGTT | AAACCTGCTG | CTACTGCTGC | TTGGCTTGCC | 6000 |
| GCACTACCTG | TACTTGGTAA | CTGGGCTTTA | TTAGTTTGAC | TAGCTTCACT | TGAATCAATT | 6060 |
| GGTTTTGTAT | CTGCTTTTTC | TGACACTTGT | GGTTTTTTAG | CTTCTTGAGC | TACTGGTTTG | 6120 |
| GTTCCAACCA | AGACGATGCG | GTCTGTCGGA | ACTTCTACCA | CTTCACGGAG | TTTTTCTTCC | 6180 |
| TTACTTCCAT | CAGGATTAAT | CGCTGTAAAG | ATACGTTCTT | TTCCAAC TTT | TCCTTCTTGT | 6240 |
| TCTACACGAG | TTTCACCTAG | ATACAGTGTT | GAATCTTTT | TCTCAACTGT | CTTGATATGCC | 6300 |
| AAATCTTTTT | CAACAAATTC | GATTTTTTGA | AGATCTTCTT | GTACAGCAGC | AACTGTCTTC | 6360 |
| TCAGAAACTG | GTTTTTCCTT | AGTCAAGTGG | ATACGGTATT | CCTTGACTTG | TTTTCCACTT | 6420 |
| TCTGAAACGA | GGCGAACAAG | TACTGAAAAG | CTATCTTCTC | CACTATCTAC | CACAGTTGAA | 6480 |
| GCTACTTGAT | TGTTTTCTTC | AACTGAGACT | TTGGCCGTT | GACCTTTATA | GGTAATTGGA | 6540 |
| TAGTCTTGAC | GATTTTCAGC | GAAATCAGCA | AGTCTTTTTC | CATCTACAAG | AATCTTTGAT | 6600 |
| TGAGTGCTTT | CTTGAGGCAA | TTCACTTGGT | GCAAGGAAGG | TCATCTCAAT | CATCGCAACA | 6660 |
| CCGCTCTTAT | CTGCTTTACG | CTCCATACGC | CATCTCATAG | CTTTGGCTTT | GATAGCTTTA | 6720 |
| AATGTTACGT | TGATTTTCATC | ACCAGCTGCA | ATGTCTTTAT | CCGCACGATA | AGGAACAGCT | 6780 |
| TCCCAATTTT | CTGGATTGTT | GAATGGATGG | TCTGCGTCGT | AGGCTTGGTA | GTTTGAATAG | 6840 |
| TAGGTTGGCA | CTTCAAAC TC | TGGACCGACA | TAGCGTTCTA | AAACGAGTTT | AGATGGTGCA | 6900 |
| TCCGTACCAC | TATCTGCAAA | GAACTGAACT | TTTCTTGTG | TAACAGTCCG | TTCTACAATC | 6960 |
| TTACCATTTT | CACGGAAAAT | CACACCCGCT | GATACTTCTG | GATTAGAAGA | TGGTGTGGT | 7020 |
| GACCAGTTTG | TCCAACGACG | ATTTTCTGAA | TGATCTCCGT | CATTGAGATA | GTCAACGCGG | 7080 |

935

| | |
|--|------|
| TCATGAGAGT TTTTGTCAAT ATCATTTGGTT GCTGAAGCAA AGGCCTGGTT ACTGTTTTCA | 7140 |
| TCATAGTTAG GGTATCTGA AAGAGTCTCA CCAAGTTTGT CTGTCACTCG TACAGTGATC | 7200 |
| TCAGCAACAA GGTTACTACC AAGGACACGG CCTCGAACAG TAAATTGACC TGCTTTTGTC | 7260 |
| AGATTTTCCG CTGGAAC TTCCTTCA ACTGTCAGGT CTTTGTTC GTAGCCGTCT | 7320 |
| TTACCTGTGA AGTAACTGG AACCTTAGTC GGCAATTCAA GTGCTTGACC TACTTGTAGC | 7380 |
| AAGCGAGCTT GTTAAACCGC AGCAACTGGT TTATGAGAAA GTAAGCTCTT ATCCTTAGTG | 7440 |
| AAGTGCAGAC GGTATCTCC TAAGATGTCG CCATTTTCAG CTTTCGCGAT GACACGAACT | 7500 |
| GGCTCACCTT CACGAACGCT TGGAACGACG GTAGCGAGAC CATTTGTGCT AACACTTGCT | 7560 |
| GTGACTGCCG GAACTTTTCC ATCTACAGAC TCAAGGTAGT AGTCTGTCAA ATCAGGGTTG | 7620 |
| AAGTTTGCTA AGTCTTTGCC GTCAACTTGG ATTCTTGTTC GTCCTTGCTT GGCTGCCGCA | 7680 |
| ACTTGTTCG CAAAGATTG TACCTCTGTG ATAGACGTT CACGCTTGT ATCTGCTTTA | 7740 |
| ACCATGCGAA TACGAACAGC ATAGGTTTCA ACTTTATCAA AGCTAAAGTG GTTCATTCT | 7800 |
| CCAGCCTTGA GTTGAGCAGG GGCTTTTAGA TTAGTAACTG GTTCCAGTT GGCAGAATCA | 7860 |
| TTAAAGACAT GGTCTCATT ACCAACAAA CTAGGGTTTT TAGGAGCTGT TGGGACAGTC | 7920 |
| TTACCAACAT AATACTCAAT CACATAAGAC TTCGGTACAC CAACTCCATG GTCTTCATGG | 7980 |
| AATCCGACAC TTAGATTATC AACGGAGCGT TTGCTCAAGA TACCTGAATC TCCAAACAGA | 8040 |
| ACACCGACTG AAGCTTCTGG ATTAGTACGA TTCCAGTTG TCCAACGATT GGCTGGTTGG | 8100 |
| TTATTGTAGG AAATGAGCTT GTCATTAAAC TTTGAAACTG GGTGCTTGG ATTTGAGTCT | 8160 |
| GAAGCAAAGG CAAGTGGCAA TTCTGAACCG GTCCATTGGT CAGAAATGTT TGCACCTTGC | 8220 |
| TCAGTTTGAG CAGATACGCG AACATGAAGT TTAGTTGTTA ATTGCGTACC TTCTAAGCGA | 8280 |
| CCATTAACTG TAAAGACACC TTCCTTAGCG TATTGCTCTG GACGAATCGC ATCCCATGCA | 8340 |
| ACCTTAGCTG ATGAAACGTG ACCATTTGAA TCATATGTCC GAACACTTTC TGGTAATTGT | 8400 |
| GGTGCTTCTG CGATTGGAGT TGTCACTAGT ACTTCTTCAA CTGAAACGAT ACCTTCTACA | 8460 |
| GAGACTTTTG CACGCGCTTC AAGGTCAATT CCTTCAACTT TACCTAGTAC TTCAAATGTT | 8520 |
| TGATAGGAGT CTAGTTTTTC TTTCGGAATA GCTTGCCAAG TGACTTTATG AGTTTTAGGG | 8580 |
| AAACCTTTGT CATACTCAAC TGTTACTGTT GCTGGAAGAC TTGGTTCCTG ATGCAAATCT | 8640 |
| GTCACACAT TTACAGGACG GATGGATTGC GCAATCTTCT TCTCAGTATT GGCTTGATA | 8700 |
| GTGAGTTCAA CTTGGTCTTT AGCTCCCTCA TATTCAGCGT TCAGAGTGAC TGCTCCTGGC | 8760 |
| TTATGCAACT CAAGCATTC TTTACGAATT GCGACTTCCC CTTCACTACT TGTAGAGAAG | 8820 |

936

| | |
|--|-------|
| GTTACTTTAT CAGCTGGTAA TACAGCTTGC GTTCCATCTT GATAGTGAGC TCGAACCGAC | 8880 |
| AATTTGACAG TTTGGTCTTC TTTGAGACTG TCAGCTTTTT CCACTTGCAA GCTCAAGTGA | 8940 |
| GCAATTTTTG GCGCTTCTTC AAGGAATTGA ATTGCATAGG TTTGAAGAGG GCCACCATCT | 9000 |
| TTAGGCTGAA TAAAGATGCT CGCACGCATG CCGTTTGCTG CGCTTGCTTG AAGAACTGTA | 9060 |
| ACAGCTGCAT TTTTAGCACT TGCTGTGACT TCTGGCAACT TAGCTCCATA AGCAAGAGTG | 9120 |
| CGGTATTGCA TTGGTTTTTG ACTAGTAAGA CCTGTTACTG CCTCACCACC AACCGTTACA | 9180 |
| GTTGGTACTG CAGGTGCCGC AGGATTGCCT TCTTCTACCA CAAGGGTTGC ATGAATTGGT | 9240 |
| TGACCTTCTA AATAACCGGT CGCTTGAATA CGAGAACCCTG GAATTGCTAA CTTAGCTTTA | 9300 |
| TCTTCTTCGG CAATCTCCCA CTTGTCCACT TCATACTCTT CAACACTTCC ATCAATCAAA | 9360 |
| ACATAGGAAA CAGATTTGTC TACAGAATTC AAGTCAGTAT TTGGAGCAAT ACGTTTCACA | 9420 |
| ACTGGTAGCT CTGATTAAAG AGCAATCACT TCTACACGAG CTTCTACTTC TCGTCCGTCA | 9480 |
| GCCATACCTT TCACCGTTAC AATACCAGGC TTGCTCACAT CTACTGAAGA CCAGGTTACA | 9540 |
| GGACGTTCTG CACGGCTACC ATCACTGTAT ACAAACGGAA CAGTGGTAGG CATTTCAGGT | 9600 |
| GCCTCTCCAA TAATGGTCTG TACTTTTGGC ACTTCTGTCC CAAAACAGT CTTCTCTGT | 9660 |
| CCTTCTTTCT TACCAGTAAA GACAGTGACT TGGTTCGATT TCAAGAGATC AGAGTGGGCA | 9720 |
| GTCAGGGTGA ATTTCCCTGC TTGTTCAAGT GATTTGACAA TGGCAACACC TTTACCATTA | 9780 |
| AATGCTTTAC GAATCCAAGA ACCATCTGCT TCGCCTTAT AGCGTTCACG GCTGGCTTGT | 9840 |
| TCTCCGTAT CTACACCGAC CAGTTGACCT TGGCCATGCA ATTGGAAGCG AACCAGATTA | 9900 |
| TTAGCAGTTG GAACCACATT CCCCTGGCTG TCAACAATTT CATAGTAGAT GTAAGTCAAG | 9960 |
| TCTTTTCCAT CTGCTGCAAT CGCATGGTCT TCCTTAATAA GACGAACTGC CGCTGGCTTA | 10020 |
| CCAGCAGTCG TAATCTTATC TCGAGCAATT TCCTTGCCAG ATTCATCACG AGCAATTGCT | 10080 |
| TCCAAGGTAC CTGGTTGATA GGCAACTTTC CATTCAAGAT AAAGTTCATT AGCATTTGCA | 10140 |
| CCTTCTTGGT AAGTCCGCCC ATCGCTGGTT TGTTTTTTAT TGAAAGTCTT AAGACCAAGA | 10200 |
| GATTTTCCAT TCAAGAACAA TTCTACACTA GAAGCATTCG AATAAGCACG AACTGGAATC | 10260 |
| TTACCTTCTG AGTCAGCTAC TTTGGATGCT AATTCTTTGT TTTCCAGTT CCAGTGAGGA | 10320 |
| AGAAGGTGTA CCATCGGTTT CTTCTTAACA GAAACCCATT GGCTTTGGTA GAGATAGAAG | 10380 |
| TCATGTTTTG GAATGCCGGC TGTATCTACG ATACCAAAGT AAGAGCTCTT AACAGGAGTT | 10440 |
| TGATTTTGGT TGTGCCATGG TGTAGGTTCA CCAATATAGT CCGTACCTGT CCAGATAAAC | 10500 |
| TGTCCAGCAT AGCCAGCGTT GTCACGGTCA AAAGTCCATG AAGCGGTTGC TGTTTTCCCC | 10560 |
| CAACCCACAC GATCATTTCC ATAATCTGAC TGTTCATAAT TACGCTCAGG TCCATTGCTA | 10620 |

937

| | |
|--|-------|
| TGTTTCAATT CACGTTTCAGG GCGATAGTAA CTTCCACGTG TACGGGTAGC TGAAGATGTT | 10680 |
| TCTGATCCAT AAATCAACCA TTTTGGATGC TTAGCTCTAA GGGCTTTGTA ATTATCTTCA | 10740 |
| GAATAGTTAA ATCCAACAGC ATCGAGTTCA TCAGCAATTT TCTCATGCCC TCCGCTACCA | 10800 |
| TTACCGAAAC GGAATTTATC TGCTCCCATG GTAACATAGC GAGTCTTATC AACATCCTTG | 10860 |
| ATAACCTTAA CCAAACGTTT AACAGTTGCT AAAGAGTGGG CATCACCATT AGCTTCACCT | 10920 |
| ATTTCATTAC CAATTGACCA CATGAAGATA GCAGGGTTGT TTTTGCCTCT TTCGACCATG | 10980 |
| GTACGTAGGT CAAAATCAGA CCATTTTTCCT CTTTTCGAG CTTCTGGGTG AGTGGCATCT | 11040 |
| TTTTCAAAGA AACGTCCATA GTCATAAGGT TTCTTGCCAC CATACCACGT ATCAAAGGCC | 11100 |
| TCTTCCTGAA CGAGTAAACC TAGTTCTGCT GCGATTTGCA AGGTTTGCTC ACTAGCAGGG | 11160 |
| TTGTGGGTTG TACGGATGGA GTTAACTCCC ATCTCCTTCA TTTGTTTGAG ACGGCGATAT | 11220 |
| TCTGCTTTAT AGTTTCTTTC TGCTCCAAGC GCCCCATGGT CGTGGTGCAA GGATACTCCA | 11280 |
| TGGAATTTAA TACGTTTCACC ATTCAAAGAG AAACCTTCAT TTGGAGTCCA GTGATAGTAA | 11340 |
| CGGTAACCAA ACAAAATCCTT CTTAGCATCA ACCAATTGAC CGTCACGGTA AACACGCGTA | 11400 |
| ATCAATTCGT ACAAGGCAGG TTTGTCAATT AAAACAGTCC AGAGTTTGG TCTTTCAACT | 11460 |
| TCTAAATCG CATCTAGGCT TGTGATTCA TGTGCTTTTA AGGTACGACT CGCTGTACGA | 11520 |
| ACTAAGCCTG TTACAGCATG ACCACCTCGT TCAACGATTT GATATTCGGC TACAAGTTCA | 11580 |
| TGGTCTTTGT CGTCCGTATT GACGATTTTG CTGGTCACAT GAGTTTCAAC CTTGCCATGT | 11640 |
| TGTTGTCTT CAAGTTTGG TGTTAAATA GTTGTCCTAT TTTTCTCAAC ATGCACCTTA | 11700 |
| TCTGTCACTT GTAAAGTCAC ATCACGATAG ATACCACTTC CTGAATACCA ACGGCTACTT | 11760 |
| GGCTGTTTGT TGAATGCATG GACAGCAATC ACATTTCTCAC GACCATCTTT TTGAAGGTAT | 11820 |
| TTGGTGATAT CATATGAGAA CTGGTTATAA CCATTTGGAT AATGCCCCAC TAACTGACCA | 11880 |
| TTGACATAAA CTTGAGAATC CATGTAGACG CCATCAAAAG TAAGGCGAAC ATTTTCTTTG | 11940 |
| AGGTCTTTTT CATCTAGTTT GAAAGTCTTG CGATACCAAG CTTCCCCACC GTTGAGCTGT | 12000 |
| CCACCTTCAT TTTGTGCAGG AGATTTCATG TCGAAATCGT TAAAGATACT CCAGTCATAC | 12060 |
| GGTAAATCTA ATTTTTCCTA CGTAGATACG TCTGCATCAG GTTTAATGGC TTCCTTAGAA | 12120 |
| TTTGCATTGA GTTTAAAGTA CCAATTTTGA TTAAATCCA CTTTCTCTGC TTCAATCATT | 12180 |
| TGATTCACCT CTTTATTTGT TACAGCTTTA GCATCTTCCT TGAGCGGTTT TTCTTGATTT | 12240 |
| GAAGCTTGTG ATTCTATCCT TGGAGCTTTT TCTTCCGTT TAGCAGACAC TTTTCTCTCT | 12300 |
| TTTGGAGTTA CGGCTTCATC TTCTTTCTTC TCAGATGCAA TAGCCTCAGT TGAACAGGT | 12360 |

938

| | |
|--|-------|
| TCACCTTTGTT CTGTCCTTTC AACTATATTT TTAGTTTCCA AAGCTTTATC AGCCTTTTCT | 12420 |
| TCTACTATCA TTTTTCCTC TTTAGGTTTC TCAGCAGTAT GAGTAATAAG TGTTCATCC | 12480 |
| GCATAAACTA CAGATTCTCC AGCTATATTT CCTCCTAATA AAAGTGCACA AGTCCCAATC | 12540 |
| ATTACTGAGC AAGCTCCAC AGCAAACCTA CGAATGCTAT AAAGTCTTTT CCGATTCCAA | 12600 |
| TGGCCTTTCC CCATAAAACC CTCCTTATAT TATATTTAGT GCAGTTAGCT ACTACCAAAG | 12660 |
| CCCAAGTGGT ATACATGGTA TGACAACCTA GTTTCACAA TTTACACTCT GCGAAAATCC | 12720 |
| AATTCAAACT TCGTCAGTGT CGCCTTGCCG TAGATATGAT TACTGACTTC GTCAGTTTCA | 12780 |
| TCTACAACCT CAAAACCATG TTTTGAGCTG ACTTCGTCAG TTTTATCTAC AACCTCAAAA | 12840 |
| CCATGTTTTG AGCTGACTTC GTCAGTTTCA TCTACAACCT CAAAACCATG TTTTGAGCTG | 12900 |
| ACTTCGTCAG TCTTATCTAC AACCTCAAAA CTGTGTTTTG AGCAACCTGC GGCTAGCTTC | 12960 |
| CTAGTTTGCT CTTTGATTTT CATTGAGTTT ATATTTTATA GGAGCGCATT ATTTTGCTTT | 13020 |
| TGCTGCGTAC TCTTCGTTAC GTTTGATCAT TTGTTTTCTG TACCAAGCAA AGATACCGAT | 13080 |
| ATAGAATACA AGGAAGACTA CTGCACCAAG GATTGCTTTG ATATCACCAG TTGTAGTGT | 13140 |
| ACCAATTGTC CAACCAAGAA GTTTTTCGAT TGGTCCTTCA AGAGTAGAGT GAGTAATCAA | 13200 |
| TTGAGTTTGG CTCACACCTT CTGGGAAGGC ACCTACACCT TTAGCAAGTT CTGTTGCAAA | 13260 |
| TGGTGCAATA AGTGACCTG AAAGAAGGAA GAGTGGCAAC AAGAGTGTTT CGAAGATAAT | 13320 |
| CATACGGAGC AATTTACCAC GAGTTACAAC CAAGAGAGCT GGAGTAACAC CCATAGCGAT | 13380 |
| GATACCTGCA AGTGGCAAGA TACCATTTCC AACTTTTGAA AGAAGCACTG CTTCAATCAA | 13440 |
| CATGATTGGT GCAAGTACGT TGGCACAAGC CCAGATTTCA GCACGACCAG CGATGAATGG | 13500 |
| CCAGTCAAGA CCGATATGTA ATTTACGTCC TTGAAGACGT TTAGTAGCAA CGTTTGTAAT | 13560 |
| ACCTTGATGAT AGTGGTTCTA CGGCTGCGAT GAACCATGAA CCGATAAGTG AGAAGAGTTC | 13620 |
| CAAAGATACA CCGGCAGTCA AACCAAGAGA CAACCATCCT TTGATAACAA GACGCCATTT | 13680 |
| ATCTGCATCT GCAACACCTG CAATTGGATG TGGAGTTCCC ATAATACCGA TAACGATACC | 13740 |
| AAGGATGAAA CCGATGAAGA ATTTAGATCC CCAGAAACCG ATTTTCTTGT TCAATTTAGC | 13800 |
| AGCATCAAAG TCATATTTAT CAAGGCCTGG GAAGAATTTT TCAAAAATCT TATCCAAAAC | 13860 |
| CATGATAACT GGGTTCATCA TGTAAGTTCAT GTGAGTTGAT GTCATTGGTG ATGAAGTTGG | 13920 |
| GGCGTTAAGA AGGTCATCAA ATGTAGGTTT CATCAAGTCA GAGTTGATAA TTTTCAACAC | 13980 |
| ACCGACAAGG ACGATAGCTG CTGTAGCAAT AAAGAGTGAA ACCCCTTGAC TCACACCATT | 14040 |
| GTTATCAGCA TACCATTTAA TCAAGAGACC TGTGATAGAC AAGTGCCAGA TATCAAAGAT | 14100 |
| ATCGACATCA AGTGATCTG TTTTCTTCAT AGCTAGCATC ACTATGTTGA CAATCAACAT | 14160 |

939

| | |
|---|-------|
| GATGAGCAAG AAGTATAGTG TCCAAGCAGA ACCCCAAGTG ATTGTAGCAA GTGGTGCCCA | 14220 |
| ACCAACGTCG GTAATACTCA ATTGGATACC AGTGTMTTCA ACGAATTTTG CTAGTGATGC | 14280 |
| TGAGAAAGCA GTGTTTAGCA TACCGATGAT AGCACCGATA CCTGTAAGAG CGATGGCAAG | 14340 |
| TTTGATACCA CCTTCAAGCG CTTTGGAGAA TTTCACTCCA AAAAGTAAAG CCAATACTGT | 14400 |
| CAAAATGATT AACATGATGA CAGGTCCACC CATTTCTAAG ATGGGATTGA AAACCTTTCC | 14460 |
| GATTAGGTCA AAGATTGCAT CCATAACAGT TCCTCCCTTT TTGATGTTAT ATGAATGTTA | 14520 |
| ACAAATTAGA ATTAGCTTAA TCCGTGTTCT TTAATAGCTG CTTCAATATT GTCAAATACT | 14580 |
| GGAGCGCTCA TTGCTGGGAT ACGGAATAAG ATTGGCCCAG CTTGATAAC TGGGATACCT | 14640 |
| GGTTCAAAAC CAAGGTCTGT TGCAGCGATT GGTGTAAAGA TATCGTAACC TTTCATAAGG | 14700 |
| TCTTCGTTTA CATCTTTCAC CATGACTGCA TCACAGTGAA CATCATAACC ACGGTTTGAA | 14760 |
| AGTTCTTCTT CTAGAGCACT TTTAATTTGG TGACTTGAGT TAACACCTGC ACCGCAGGCA | 14820 |
| GCAAGAATTT TAATCATTTA GATTTCTCC GATTTTATTT TTTAATAGAC AAGATTAAGC | 14880 |
| GGTTGCTTCA GCAATGTAAG TATAAAGGC TTCTGGTTCA GAAATTTTGT ATAGGTCTTC | 14940 |
| AAGATGACCA TTTCCTGTGA AGAAGTCCAT TAACTGAGCA AGAATGTTCTG TTGACTTGA | 15000 |
| ACTTGAATTA TTAATGATAA AGAAGAGTAG GGATACTTCT ACTTCCTTAT CAGGAGCTAT | 15060 |
| CATATTGTGA AAAGTTATTG GTTTTCTAA TCGAACAACC ACCACTTTCT CAGCTAGATT | 15120 |
| ATGAACAATA TCTGTGTGAG GAATCGCTAC ATTTGGCAAG TCCTTTCCTA GAAATTCAT | 15180 |
| ATCTAAACCA GTTGAAATG ACTTTTCACG CGTGATCAAG GCTTCACGAT AAGTTGGAGT | 15240 |
| GACAATTTCT CGTTCTTCCA ATAAAGTGC AACCTGATCA AAGAGTTGTT CTTGACTATC | 15300 |
| CGCTTCTAAG CAAACACAA GGTTTTGTG AAAGAAATAA TCTAATACCA TAAGTTTTC | 15360 |
| CGG | 15363 |

(2) INFORMATION FOR SEQ ID NO: 140:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 28882 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 140:

| | |
|---|-----|
| TAAGACTATT TAATAGTGGA GTGAAATAGG ATACGAACAA ATTGATTAGG AAAATCAAAT | 60 |
| GAATTTATAG AAATCTTTTA GCAGTTATGT TATCCTATTC TAGTTTCAAA ACGCTATAGA | 120 |

940

| | |
|---|------|
| AGCAGCATTG TGCTAGTCKA GATTCAGTTT ACTATACTAA AACGAGTAGC TTGAAATCAA | 180 |
| AAAACCCACC CTCACAGGCA GGTTTTATCT GTATTATTCA GCTAGATTAT GCTTTACCTT | 240 |
| CTGAACCGAA TACGTCGATA CGTTCCTCAA CCGATGCTTG GATAGCTTTT ACACCGTCAG | 300 |
| CCAAGAATTT ACGTGGGTCG AAGAGTTTTT TCTTGTCGTA TTCTGCTTCG TTTGCTTCGT | 360 |
| AGTCACGAGC AAATTTACGA GTTGCCTTAG CGAATGCGAT TTGGCATTCT GTGTTAACGT | 420 |
| TAACTTTGGC AACACCAAGT TTGATAGCTG CTTGGATTG CTCATCAGGA ATACCTGATC | 480 |
| CACCGTGCAA TACGATTGGG AATCCTGGAA GAGCTTCTGT CAATTTTTCG AAGTGGTCAA | 540 |
| GGTCAAGACC TTCCAGTTT ACTGGGTAAG GACCGTGGAT GTTACCGATA CCAGCTGCCA | 600 |
| AGAAGTCGAT ACCAGTTTCA ACCATTGCTT TAGCGTCTTC GATTGGAGCC AATTCACCTT | 660 |
| TACCGATGAT TCCATCTTCT TCACCACCGA TAGTACCAAC TTCAGCTTCT ACTGAGATAC | 720 |
| CTTTAGCGTG TGCTTTTTC ACAACTTCTT TAGCCAATTT AAGGTTTTCT TCAACTGGAA | 780 |
| GGTGTGAACC GTCAAACATG ATTGAAGTAT AACCAACTTC GATACACTCA AGTGCACTCT | 840 |
| CGTAGTGACC GTGGTCAAGG TGGATAGCTA CTGGTACAGT GATACCCATT GATTCAACAA | 900 |
| GGTTAGCGAT CAAGTTGCGA GCAACTTTGT AACCACCCAT GTATTTAGCA GCACCCATTG | 960 |
| AAGTTTGAT CAAAACGTGA GCTTTTTCG CTCTGCTGC GCGCAAGATA GCTTGAGTCC | 1020 |
| ACTCAAGGTT GTTGTGTTA AATCCACCAA CTGCATAACC GTTGTACAGG GCTGCTTGA | 1080 |
| CAAATTTTTC TGCTGAAACG ATTGCCATTT TATCAGGCCT CCTGTATATT TTTATGGGTC | 1140 |
| ATCCCATTTA CATTTGTCAT TTTATCACTT TTTGCCAAAA AAATCTAGTT TTTCCCGCAG | 1200 |
| TTTCGATTGA TTTCTTCTA ACTCCATCTA TGTAACCCCT TTCTCTCCCT AGTCTTGGAC | 1260 |
| GACTTTTGA AAATCTATAA AGAAGGTTAA ACTATTCTCC TCCATCTCGA AACGATAAGC | 1320 |
| TAATTTTTC TGTCTAATA GACTCTTAAC CACAAAGAGC CCCATACCAG ACCCCTTGAC | 1380 |
| CTTGCGACTG GCATTGTCAG AAAAAGACTG GGCTAGTTTT TCTTGTTCCT CTGAGCTACA | 1440 |
| GCTATTTTCG ATAAAAAGTT CTCCTTCTCT TTCTCCAATT CGAACTAAGC CACCTGGAAC | 1500 |
| AGAGTGCTTA ATGGCATTGC TGATGAGATT AGAAAGAATC AACTTCATAA CTGATGGGTT | 1560 |
| TAGATAAGCC TGCTGATGGG TCAAACATTT GTCTATCTGG AGCTCTCTTT CCTTGGCTAG | 1620 |
| CAAGGCATAA TCTTTGACCA GATTTTGCGT CATCTGGAGG AGGTCAATTG TTTCCCTATC | 1680 |
| ATCTCGCAAT TCCTGCACAG AAGAGAGGGA AAGTATCTGC AGAACATGGT GATTGAGTTC | 1740 |
| ATCCACAATC CCCAAGGCAA CTCCAGATA CTGGTCTCTA TCCTTATAAC GACCGATATT | 1800 |
| CTCTCTCATA TTTTCGATTA GGATTTTCAA ACTAGCCAGC GGTGTTTTCA ATTCATGAGA | 1860 |
| AGCTCCTCGT AGGAATTCGA CCTTCATCTT CTCCAGCTGG AGAATGGCTT CATTCCTTTC | 1920 |

941

| | |
|--|------|
| ATGCAAGTCC GCAATAACAG TCAAGAGATG CTGGTAGAGG CTATTGATTT GTTCCTTGAG | 1980 |
| ATTACCTATC TCATCCTTAG AATCCACGCG CAATCGCACT TGGGAATCCA GTTCCATCAT | 2040 |
| CCGACGGGTC ACCCGCTTGA TTTCCAAAAT CGGTGCAACA ATAGTCCGAG CGTAGATGTA | 2100 |
| GGCCACCAAA AGGGAAATCA GAAAGGAGGC CAGCAAGGTA TAGGGAAGAA ACTGGAGACT | 2160 |
| GATTTGCTCC GCTTCCTTTT GTAAATCCAT GGAAGCTAGA AACTGGAGAA TCATAGTACC | 2220 |
| ACCGTCTTGC GTTTTCACCT CGCGCTCCTC AATAAGAGA GAGGTGTCTT GCGGCTCTGT | 2280 |
| GTCCAGAGGA AGACTGTCTT TGAATCTTAA CTGTCTCTCG GTCATCTCAC CTTTGACGGT | 2340 |
| CCCCTTGATA TCACTAGTCT GGAATACAA GTCTAACACT TGCTCGATAC TCTGCCTATC | 2400 |
| TTTCCCTTCT AGGGACTGGG CAATGGCTGT TGCCTTTTGA CCAATGGTTT CCTGACGATG | 2460 |
| ACTCAGATAA GTCGAAGGAA AAAGAAAATA AATAGCTAAA TGAAGGCAGA TAACCAGAAC | 2520 |
| ACTAAATATC GAGAAGGTAT AGATAAATAT CTTTGCAAAT AAACCTGTTC GTTTCATTTT | 2580 |
| CGCTCCAATT TATAACCAAC ATTGCGCACA GTGAGGATAC AATCCAAGTC TAGCTTTTTC | 2640 |
| CGCAATTCCT TGATATAAAC ATCAATAACA CGGTCAAAGG GAACCTCATC TGTCGCTTTC | 2700 |
| CAGACGGCAT CGATAATCTG AGATCGAGTC AAGGCCCGGC CTTTATTTT CACTAGATAG | 2760 |
| TCCAGAATTT CCAACTCTTT GGCATTGATA GGCACCTTCT GACCTGCGAG GCTTGCACTG | 2820 |
| TAGCTTTCAA AGTCCACCTT GGTATCCTTG TAAGAAAAGA TTCGTCTCTG ATCGTAGTAG | 2880 |
| CGCTTGAAAA TCGCGTCCAC CCTCACTTTT AAAAGGAGA GGGAGAAAGG TTTTTCAGA | 2940 |
| TAGCCATCTG CCAAAGAGGC AAAGGCACTC ATCTTGATTT CCTCATCTTG AAAAGCTGTC | 3000 |
| AACATCAAGA CAGGAACCTG ACTGGTTTTA CGAATCTCAG CTAGGACTTC TAAGCCGTTG | 3060 |
| AGCTTGGGCA TCTGGATATC CAGTAAACC AGGGCCACCT CATAGCTAGA AAATTGCTCC | 3120 |
| AGAGCTTCCT GACCGTCCGC TGCCTCAATA GTTTCATAGC CACAATCCGT CAAATAATCA | 3180 |
| CTGACCCCTT CACGGATCAT CTCTTCATCT TCTACAATTA AAATTTTCAT ACTTTAACTG | 3240 |
| CTCTCTATTT TTTATTTTTC TTAGAATAAA TACCTACCCT ATTTTCTATT ATAGTCTCTT | 3300 |
| GCTGGCCTTT TGTCTGCAAG CAACTGACCA CTAGATAAAA CGTTGTGAAA TTCCTTTCTC | 3360 |
| ATAAATTCCTA TAACTTTAGT ATATTATATT TAAGCACTAA AGTACAAAGA AAGCAACTGA | 3420 |
| AAGCAATGAT TTTCACCACT GCTTTCGGAT TTATTTTGAA TTGTTAAATA GCCATTCTTA | 3480 |
| TCCACTATTC TTGAATAGAA ACACAAGATG CAATCTTTAT TCTAGACTCA TTTTTCAAA | 3540 |
| TTTATTCAAC ATCCAGCAAG AGCTCTTTTG GTTGTCTTCT AAGGAGATTG CTTGAAGCAA | 3600 |
| GCGCCATAAC GAGAACCCT AGAACCAAGG CAAGGACAAA AATGATGATA AAGTCTGATG | 3660 |

942

| | |
|--|------|
| TCTGAATGGA AATGTCTAGG CTCGACAAGG TCTTGCTAAA GCCATCTACT TCTGCACCAC | 3720 |
| CACCAAGGTT AGAGGCTTGA GCCGCCCTTAC TAGCCTGTTT GGCAACACCT GAAGTCACAT | 3780 |
| TGGCAAGGAC AGTGTTCCTCA ATTGCACGGG CAGTGTAATT AGCTAGGAAG TAAGCAGAAA | 3840 |
| CTAGAGCAGG GATAGCAATC AAGATAGATT CGGTGATGAA TTGACCCAAG ATACTTGCCT | 3900 |
| GCTTGAGGCC GATAGAGAGG AGAATTCCCA CTTCTTGCG ACGGGCGTTG ATCCAAAGGC | 3960 |
| TGAGCAAGAG GGCAAGGAGG AGAACTGAGA AGTCAAGCT ACCCCAGAAG AGGAGGTTGG | 4020 |
| CCATCTTGTA CATACCAGAG ATAGATTGCT CAAGAGCTGG GTAGTTAGAG GAGCTCTGA | 4080 |
| CGAGTGCTGA GCTCTTCCAG TTGATACCAC TGATGCCATT CAACTCTTTC ATAACATCAT | 4140 |
| CCAAGTTCTT GTCTGCTGTT ACAAAGAAGG TTGCGTCCCC ATAAATGGCT GTGTCTTCTG | 4200 |
| TGTATCCATA AAGTTTTCGA GCAGTGTGAA TGTCTGTAAT AGCTGTGTTT TCGTAAAGTT | 4260 |
| CTTGAGTA GGTACTGCT GACTTATTAT GACCATCAAA GAGTCCCTTG ATTGTCACCT | 4320 |
| CAACTGTTC CTTGGCTCCT TTTTATTAT CTGCATCGTA GATATTAGAG TCCAGTTTAA | 4380 |
| CCTGTCCCC TACTTCCAG CCGTGTGTTG CTGCCAAGTC CTTGTGCAAG AGGATTTTAT | 4440 |
| CCTGTGCTC GTTGGTTAAG TGCTCTCCTT CGACTAGTTT ATAAGAACCA GAGACAACT | 4500 |
| TGTCTTCTT AGAGGAGTCA TTGACACCTG TAATCATCAA GCTACTTCCA AAACGCTTGG | 4560 |
| CACGATCAGC AGTGAGATTC TTCTTGGTTT CTGGCGTTTC AATCAGGTCA TATCCAGTCA | 4620 |
| AATCTCCGAT AGCGTTGATA CGTTTGACAT AAGACTCAAT GGCCTTGTTT TCGGTGATTT | 4680 |
| TTTGTATGTC TTCACCCCTG ATATTCCCAG CACCACGAGG CGTTCCTTGG TTGACGCGAC | 4740 |
| GATTGATTG CATGGAGAAG CTATTGGTGA TATTTTAAA GGTCTCCTGA GAAGCCTTGG | 4800 |
| CAGTAGCTCC CTTGATTGAC AAGCCGACCA AACTCAAGCT CGCCATGAGG AGAATAATCA | 4860 |
| GGAAGATGAC AATCGATTTG AAAAATTCC TTGTAACATA GGCAAATGCG TTGTGTAACA | 4920 |
| TAGATTCCCT TTCTAGATTT TGTTTAAATC ATTCTATTAA AATAAGCTCA AATTATTAC | 4980 |
| TAGTATTGCG CGTTTCAGTC AGTTTCTTAT CCTTTAATTC AAGTGTAAATA TCTGACGCTT | 5040 |
| GTGCCACTTC TTTACTGTGA GTTACGACAA TCACACATTT ACCTGTTTTC TGGGCAAGTG | 5100 |
| ATTTGAGTAG TTCGACAATA TCTCCAGCAG TTTTAGGATC CAGATTTCTT GTTGGCTCAT | 5160 |
| CAGCTAGAAT AACTGGAGCT TCTGAGACCA AACTGCGAGC AATGGCAACA CGTTGCTGTT | 5220 |
| GACCACCTGA TAACTGGAGA ACATTCCGCT TGATCTGGCT TTCATCCAAA CCAAGCTCAA | 5280 |
| GAAGTGTATT CTGCTTGCC TTTTGTGTA CCAATCGGAT ATTTTCCAGC GGAGAAAGAT | 5340 |
| AATCTATCAA GTTATAATTT TGAAAGACCA GGGAAATATG GTGCATGCGA TGGTAAGAAT | 5400 |
| AGCCCTTCTT ACGAATATCC TCTCCTGAA AAAGGATAGA ACCTTCAACA GGAATATCTA | 5460 |

943

| | |
|--|------|
| GACCAGCAAG TAGGGACAAG AGTGTGGATT TTCCTGCTCC TGA | 5520 |
| CTCCCCA ATAATACTGT | |
| AAAAATTTCC GGGTTCAAAA TTATAATTGA TCTGATATAG GACTGCTTCA GCAGTATTCT | 5580 |
| TATAACGGTA GGTAACATCT TGTAATTGTA ATAAAGTCAT GATTTCTCCT TCTTAACTAA | 5640 |
| TAGATGATAA AATTTCTTTC GGTGATTTTC TAAATAAGAA TAGGAAACAA AGGGCTACAG | 5700 |
| ATAAGCAACT AAGCAGAACT AGAAAAACAT AGGATTCTGC AAAAGATAAG ATGCTAGTTG | 5760 |
| ATAAACTGCT TGCTTTGGCT AGTGATCTT GTAAGCTTGC CTGATCTCCA CTGCTAGTA | 5820 |
| GAGTTTGAG TAGGTAAGTT GTGATTGCGT TTCCTGCAAC AAATGCTGGA AGCAAAGCTC | 5880 |
| CAAGAGATAC CAAAACCTACC TCTAAACAGA ATTGTAGGAA GATCGAGCTC TTGCCTTTTC | 5940 |
| CAAGTGCAAG TAAAATCCCC ACTTCATAGA CCCGTTCTCT CAACCAGAGA GACAAAACCA | 6000 |
| GAATTAAGGC TCCAGCTCCT GCTATCAACA TCCCATAAAG GAAGATGGTC AGGAAGGTTT | 6060 |
| GGAAGTTGC AACTGAGTCT TTGATTTGTT CAAAAGCCTT GTTTTCCTTT TCGACTTGGT | 6120 |
| AGCCTTGATT TTCCAAGGCC AAGTTTCTA CCTGCTTCAT GAGTCCGTCC ATTTCTTAG | 6180 |
| GATTTTCTAC ATAGAAGCGT GCTGCACTGA CTTGAGCTTC ACTATTGCCC AAAAGGGTTT | 6240 |
| GGTACTTTC ATAGTCTGTA AAGACTTGAT TTTCACTGAA GTCAGAAGAC AAGCCTGTGA | 6300 |
| ATTTCTCTTG TTTTACCA GAAAAGATGC CGATAATCTC AACTCTACT GTTGTCTCT | 6360 |
| TTCCAGATTC AACTGACCA GCATCCAAGC CAATCTTGTC ATGAAGCGAA AGACCGTTCT | 6420 |
| TCTTAGCCAA TTCTTCGTGG ATAAGGATTT TCTTGAATC CCCTTTTGA AGGTGTCGCC | 6480 |
| CTTCTTTTAG ATTGAAAGCC GAACTGGTAA AGGTTACATC CTTGGATGAA TCCTCAAGAG | 6540 |
| CCGTAAAGCT AACCAAGTTA TTGTCTGCAG CTGATAAATC ATCACGCTCC ACGCTCTGCT | 6600 |
| CGCCAGTCAC TGCTTCCTTG TCTTTTAGTT TTGCGACCGT CTCAAGTTCA GGAGAGACAT | 6660 |
| TTCCAGCCC CTTAATCTTG CTTACAGATG CTAGGTCTGA CAACTTGAAT GTCTGACCAT | 6720 |
| TCTCTATCTT CTTAATAGAA AAAGATGTAT TGAGTGATTT ATAAAGATTG CTTTCTACTG | 6780 |
| TTTGTGGA CTTTATCAGA GTCAAAACAGG CTGAAATTC GGCCAATAAG ACCAATAAAA | 6840 |
| TCAGAAATAA AATAAACTT CTCAGTCGCT TTCTGCTGAC ATAAGCCCAA GATCTTTGGA | 6900 |
| TTGGATTCAT TTGTCACCTC CATATTTGTA AGACTATTAT AAAACCCAAA TATGAAATAT | 6960 |
| TTATGAAATA CGAAAAAAA ATATCGAGTA GGGGATAATC TCTAGCCCCT CTCACACCAC | 7020 |
| CATACGTGCC GTTCGGCATA CGGCGGTTCA ACTAACTTTT AACGCATGTC GTTCAAGGTA | 7080 |
| ATAATCCAAA CACGAAACCA GTCCACGTTT TTCAAGGACT GGTTTTGATA TAGCACGTTT | 7140 |
| AAGTACCGAC TTCTGAGCTA CTATAGTAGA TTGAACTAG AATAGTACAC CTCTACTTCT | 7200 |

944

| | |
|---|------|
| AAAATATTGT TAGAAATCGA TTTGACTGTC CTGAACAATT CGTCCTATTC TTATTTCATT | 7260 |
| TTACTATAAT TGATAGTGGT CGCCCCAGCC AGATACCTTA TCTGCTATCC ATTTAGGAAC | 7320 |
| CCCTAACTTA AGCAATCCCC ATAATCGTCT CGATTTCCTC TTCCATTGCT TCCAGATAAT | 7380 |
| CACTCGTAGG CGAGTACGCA AGCGCTCATC TATGCTAGTG ACTATACTTT TCATATTTAT | 7440 |
| AATTCATTCC TTTCGTTTCA CTCAAGGCAC AACACAGAAT GAAAAAGTGT TGTGATCTTT | 7500 |
| ATTTTGT TTTT ATAATAATAG TGAGAAAACC TATCACTACT ACAAATCACG GGGAGGTGAA | 7560 |
| TAAGTGAGTG GTACAGCCAC TACCTCGCAT ATTTTGTGAC ATCATTTAAC GGTACATAAT | 7620 |
| AAGTTGTACC ATCTGAATAA GTTGCTACAA TATCATTTGC ATGCTCTCCT TCACCTTTAG | 7680 |
| CAAAGGTTGG AGCTCCTGCT GGATGATTTT TATTTGCCTC TTTCAAATTT TCAATAATGG | 7740 |
| CATTTTTTCT GTATCTTTTA TATTATCAGG ATTTTTCCT AAGATTTTGT CTGGATATGT | 7800 |
| CGGTTTAGCA GAAACAATTT TTAATTTTAC TTCTTTTTTA TTCGAAGCAC TTGTCCAGTT | 7860 |
| TCCAGCATTA TCTTAGCAT TTAATTTTAC AGTAATTCCT GAACTAGGAA CTTCAGTAGC | 7920 |
| AGGTTGATTA TCAACATTAT TCAACTTTAA TTTCAAAAGA GCTGTTGCAT CAGACGTTTT | 7980 |
| ATCAATCGTT ATATATAATG ATGAATGTG ATTATAACA GTTCCTTCAT ATTTAGCTGT | 8040 |
| TTGTGAGCTA CTTGAAACAG AACTGAAAT ATACCCACTA CCTCCCTGAT TATCTTCAAT | 8100 |
| GCTTACGTCT AAATGAACTT CCCCACTATT ATTTGGCTTA GCAACAACCTG TTATAGTAAA | 8160 |
| ATAACATAAA ATTTGCATAA ATAGATTAGG GAAATCAAAG CAGCTTCTAG GAATGTTTAA | 8220 |
| GCAGTCACAG TGTACTTTCC CAGCATCAAG CCACTATAAC TCTGCACATA AAAATGGAGA | 8280 |
| AGATGGCAAT CCTCTTCTCC AAATATTAAC TTCTTTACAA ACCAACTATA GTTGACAAAG | 8340 |
| AACCTAAAAT CAATTGATAA CACAAGGTCA GGTCGGTCAA CTCTTTCAAC TGAAGCCCTG | 8400 |
| TCAACTCTTC CCATTTATCA ATCTTGATTT GGAGAGAATT GCGGTGCAGA TAGAGTTGCT | 8460 |
| GGGCTGTTTT AGTGAGAACA GCACTATTTT CCCAAAGAGA GAGAATGATT TCCTGAATCT | 8520 |
| GATCTTGATC CAAAATCATC TGGTGTAGAC ATTCCCTGAT TGGCTTCAAG TCCACGAGTC | 8580 |
| TTTCTCCCAT ACTCCAAAGA TAGAGCTGAG AAAAAGTATG AACACCTTGG TGACCCTGAC | 8640 |
| GCCACCATGT CTTGAACAAA TCCCCTCAG CTTTGATTAA GTCTGATAGG GCTTGATGTC | 8700 |
| CCGTCTGAGA CCAAACCTGA CCCAACATGA TAGAAAGACG AAGTCCAAAG TCATACTCAA | 8760 |
| CCGCTTCAAT CGTATCACTT AAAATATCTC TTACAGAAGT GTATTTGTCT TGTGAAGCA | 8820 |
| CGAAAACATA ATCCTGAGCT CCGACCTGTA GCACTGTCTG ACAATTCGGA AAAAGAGTCC | 8880 |
| GCATCATATC TAGCCAAGAA GCCAGATTTT CCTGCTGAAA ATAAGAAAGA TGGCAATAAA | 8940 |
| CCAACTGAAT CTTTTTAAAA ACTTGCGGTG CCTGTCCCTT GCCCTCAACC AGATAGGAAT | 9000 |

945

| | |
|---|-------|
| ACCAAGGGTT TAGCGAACGA GCCTGCTCCT GCTGGGTCAA AAGGGCAACC AACTGCTTTT | 9060 |
| CACGCTCGCT GAGCCCAGCT TCCTCCAGCA AAATCCACTG CTGAGAAGCT AAAGGGAGCG | 9120 |
| TGAGATAGCC CTCTTTCTCT ACTGGTTGGT CTGAAATCCG AGCCTCAGGA AACCAGTCTT | 9180 |
| GTAGTTCTTT TGCCCTCATG TTCTAGCCCT CCACCTTTTG GATGCACCAT GAAACCAAAC | 9240 |
| TCTCAAGACG TTCCAGATTG TCAGTCATAT GGAGATAGCC CATAACCGCT TCAAATCCCG | 9300 |
| TGGACATACG ATAAGTCACG ACATCTGCAT TTTTAGCCTT TGTGTGGCTA TTGGTATTGC | 9360 |
| GGCCACGTTT GTAGATTCTT TCTTCTTTT CCGTTAGGAC CTGCTCCTCC AACATGAGAG | 9420 |
| CAATCAGGCG AGCCTGAGCC TTGGCTGACA CGTACTTAGT TGCTTCTTGA TGGAGTTTAT | 9480 |
| TGGGTTTGGT CATACCTTTG AGGATGAGGT GACGGCGAAT ATACATAGAA TACACCGCAT | 9540 |
| CCCCCTCAAA GGCTAGCGCA ATCCCGTTAA TGAGATTGAC ATCAATCAGG TGTCCACCTC | 9600 |
| ACTCCATCCT TGGTATCAAG GAGCTTAATT CCTTGAGTAA CCAATTGGTC ACGGATTTGG | 9660 |
| TCTGCTGTCG CAAAGTCACG ATTGGCACGC GCCTCTTGGC GTTTTGAAT CAAGTCTTCA | 9720 |
| ATCTCTGCAT CCAAACTTC CTCAACAAAG ACAATTCCAA AAATTTCTAA CATATCTGCA | 9780 |
| AGAGCTTGCT TGACACTTGC ATCATAGTTC CCTGAGTTGA TCCATTTGGC CATTTCAAAG | 9840 |
| ACAACTGTGA TACCGTTGGC AGCATTAATA TCTTCATCCA TAGCTGCTAC AAACTTATCT | 9900 |
| TTAAAGTTT GTAACCTTG GGCATCCACA TTTCCTGTAA ATGGTTGTTC GTAAGTATTC | 9960 |
| TTGAGATACT TGAGATTGGT CTCGGCATCG CGAACTGCCT TTTCCGTGAA GTTGATAGGC | 10020 |
| TTACGGTAGT GCTGGGTCGC AAAGAAGAAA CGAAGTACTT GCCCATCAAG AGTTTAAAGG | 10080 |
| GCATCGTGTA CCGTAATGAA GTTACCCAAG GACTTAGACA TTTTGACATT GTCGATATTG | 10140 |
| ACAAAGCCAT TGTGCATCCA GTAGTTAGCA AAAGCCTTGC CTGTTTTAGC TTCAGACTGG | 10200 |
| GCAATTTTCAT TGGTGTGGTG TGGAACTCT AGGTCAGCTC CACCACCGTG GATATCAATG | 10260 |
| GTATCACCTA AAATCTCTGT CGACATGACT GAACACTCAA TATGCCAACC CGGACGTCCA | 10320 |
| GGTCCCCAAG GACTATCCCA AGAAATCTCA CCTGGTTTGG AAGATTTCCA TAGAGCAAAG | 10380 |
| TCTACAGGAT TTTCCTTACG AGCCGTTTCT TCATCGGTAC GACCTGAAGC ACCTAGCTCC | 10440 |
| AAATCTTCCA AGGTTTTATT AGCCAATTTA GCATAGTTGT GGGATTTTTT TACACGGAAA | 10500 |
| TAGACATCCC CTTGACTCTC ATAGGCAAAG CCTTTCTCGA TCAAGTCTTC CACAAAACGG | 10560 |
| ATGATGTCTG CCATAAACTC CACTACACGC GGATGGCGAG TCGCAGGTTT CACGCCCAAT | 10620 |
| GCCGTCACAT CCTCACGAAA GGCAGCGATG TACTTATCCG CAACCTCCTG AGGCGTGATA | 10680 |
| CCTTCTTCCC TGGCACGGTT GATAATCTTA TCATCCACAT CTGTAAAATT GGAAATATAG | 10740 |

946

| | |
|--|-------|
| GCAACCTTAT ACCCACGGTA CTCAAAATAG CGACGAATCG TATCAAAAGC TACCGTCGAA | 10800 |
| CGGGCGTTTC CTACGTGGAT ATAGTTGTAC ACCGTTGGCC CACAAACATA CATCTTGATC | 10860 |
| TTGCCGTCCT CAATCGGGAC AAATCTCTCGC AAATCACGAG ACATGGTGTC ATAGATTTTA | 10920 |
| ATCATAAATC ATAATCAGGA AAGCTGAAAT CCAAGAACAA TTAGTTTCAT CACTAAAAGT | 10980 |
| TCAAGTAAAT TTCAGTCCGA ATATCTCTAC ACTTCGGAAT CCCTTGCTCC TTTCTCATTC | 11040 |
| AGATAAACCA CCTGAGTCTG TTTGACAAAG CCAATTTTTT CATACAAACG TTGGGCACCT | 11100 |
| ACATTGCTAT CTTCCTCTGC AATCTGAAAT TCCTTGTCAT TTTGCTCAAT TAGTTGGTTG | 11160 |
| ACGAGGGATT TTGCTAAGTA GCTTCCATAG CCTTTTCCAC GTTCAGGTTC CAATATTGCT | 11220 |
| AAACCGTAGA GGTAATTCGT ATTAGTCGAT AAATCAACCG TACAAGTTC AATAACCTGA | 11280 |
| CCAGCTTTTA ATAAATATA TAGTCGGCTT TCTGGATCTT TCAGAGCTTC AGCGACATAT | 11340 |
| CTATCCACAA CTCTCTCGA TTCTGTTCC TCTGAAATG CCTGAAATTT TAATTGACTA | 11400 |
| ATTTGATCCT GATACGAACT ATCTGCTAAC AAACTTCAA GATGGGAAAC ATTTGCTAAC | 11460 |
| GGATAAGGTC TTCTATCCTT ACCTAACCA GTTTCTGTCT CTTCATCCTC GATTAGTCCC | 11520 |
| CAGTTACTGG CAAAGTCAGG ATGATTCTCT AAAAAATAC GTTCTGTCTG AAAAGTGA | 11580 |
| GACCGAATGG GGAAAGAAGC TGTTCCTCTC TCAAACTAG TAAACAATGC ACGCGCAATC | 11640 |
| CCCTGACGGC GATGACCTGG ATGAACCACT ATCGTCACTT CTACATCTTG GTCATCTGCA | 11700 |
| TAGACAGTTA ATAAACCAAC AAGTTCGCCT TTTTCATAAT AAAGGAAAAA GGCGGGCATG | 11760 |
| TTTGGGTCAA AATTAAGCAT GTTAGAGAGA TAGGGATCGC GATAGGTACC GTCATAGTTT | 11820 |
| TGGCAACAGT TAATTACTTT TTTGCCTCA GATAGCTCCT CTGGCTTAA CTTGTTCTTT | 11880 |
| GCTTGAATCA TATAGGTATC CTCTACAAAC CAGACGATCT GTGACTGGCA TCTTTAGCCT | 11940 |
| GCTCGAGTTT ATTGACATAA TACTCTCGTT TTTCTCGAC TTCGTGAATG ACAGGCTCAT | 12000 |
| CTTTCTTACC ATGAAGACGG ACAATCTTGG CCGGAATACC GACAACCGTC ACGTCACTAG | 12060 |
| GTACATCTGC TACGACAACT GCTGCAGCAC CGACCTTGGC ATTTTCACCA ATTTCCACAG | 12120 |
| GCCCGATAAC TTGGGCATGG GCTGATATGA GGGCTCCCTT TCGTACAGTC GGATGGCGTT | 12180 |
| TGCCACAGTC TTTCCCTGTT CCCCCGAGAG TCACTCCGTG ATAGAGAAGA ACGCCTTTT | 12240 |
| CAACAATCGC TGTCTCTCCA ATCACCAGAC CAGAACCATG GTCAATAAAA ACACCTGAAT | 12300 |
| CAATCTGGGC TCCTGGATGA ATCTCAATCT GAGTCCAAAA GCGCCAAAAC TGACTGTACA | 12360 |
| TACGAGCTAA TAGTTTGAAG CCGTGCTTCC AGAGAAAATG CGAGAGACGG TGGGCCGCCA | 12420 |
| AGGCCTTGAC ACCTGGATAA GTCAGCAAAA CCTCCAAAGT GGTGCGGGCC GCTGGATCAT | 12480 |
| TTTCTTTTAC AATATCAATG GTTTCGCGCC ACCACCCCAT ACATTTCTCC TTTCTTATT | 12540 |

| | |
|--|-------|
| CTGAATCTTT TGATGTTTCT GTAAATTCCTT TCTTAGGTTT GTAATCCTTT TGATGACGTG | 12600 |
| GGCGGTGAGG GCGCTCAGAC TTTCACCTT TTTCATCATG CTCAGGTTTT GGCGGACGAG | 12660 |
| GTAAGAGAGC CTTTCATAGAG GCATCGATAC GGCCTTTTTC ATCAATTTTG ATAACCTTAA | 12720 |
| CATCAACTTC ATCCCCGATT TCTACCAAAT CCTCTACAG ATTGGTACGA GTCCAAGCCA | 12780 |
| TCTCAGAGAT ATGAACAAGG GCATCTGTCT TATCAAAGAG GTTAACAAAG GCACCAAATT | 12840 |
| TCTCGATACG AACGACTTTA GCACGGTAAA CTTTCATCCAC TTTGGCTTCA CGAACCAAAC | 12900 |
| CAGCAATAAT TTCTTTGGCA CGGTAAATAG CATCTTGGTC ACTAGAGTAG ATAGACACAT | 12960 |
| TTCTTCTTC GTCTATATCA ATCTTAACAC CTGTTTCAGC GATAATCTG TCGATGGTTT | 13020 |
| CTCCACCTT ACCGATGACA ATCTTAATCT TGTCCACATC AATCTTGATC GTATCAATTT | 13080 |
| TCGGAGCAGT TGGAGCCAAT TCTGGACGAA CTTCTGGAAT GGTGCTTCA ATGACATCAA | 13140 |
| GGATTTCAA ACGCGCTTTC TTGGCTTGAG CAAGAGCCTC CGTCAAGATT TCTGCAGTAA | 13200 |
| TCCCTTGAAT CTTGATATCC ATTTGAAGGG CTGTAATCCC ATCAGAGTA CCTGCAACCT | 13260 |
| TGAAGTCCAT ATCTCCAAAG TGATCTTCCA AACCTTGGAT ATCTGTCAAT ACTGTGTAGT | 13320 |
| TATTTCCATC TGAGATAAGC CCCATAGCAA TACCAGCTAC TGGCGCCTTG ATTGGCACAC | 13380 |
| CACCAGCCAT AAGGGCAAGA GTTCCCGCAC AGATAGAAGC TTGAGATGAA GAACCGTTTG | 13440 |
| ATTCAAAAAC TTCTGCTACT AGACGGATAG CGTAGGGGAA TTCTTCCAAG CTTGGCAAGA | 13500 |
| CTTGAGCAAG AGCAGCTCA CCAAGGGCAC CGTGACCGAT TTCACGACGA CCTGGCGCAC | 13560 |
| CGTAACGACC TGTTTCCCCT ACAGAAATTT GAGGGAAGTT ATAGTGGTGC ATAAAGCGTT | 13620 |
| TCTTGACTC TGGATCCAAA CCATCAATGA TTTGAGTTTC TCCCATCGGA GCCAAGGTCA | 13680 |
| AGACTGAAAG AGCTTGAGTT TGCCACGAG TAAAGAGACC TGAACCATGT ACACGAGGAA | 13740 |
| GGAAGTCAAC AACCAGATCC AAAGGACGGA TTTCATCGAC CTTACGACCA TCAGGACGCA | 13800 |
| CCTTGCTTC TGTAATTAAA CGTCGCACTT CTGCGTGTTC CATTTGTTC AAGATTTCAG | 13860 |
| CCACATCACG CATAATACGG TCAAATCTT CGTGGTCCG ATATTTTCT TCGTAAACGG | 13920 |
| CAGTCACTTG GTCTTTCAT ACTTGAGTCG CAGCTTCACG GGCCAATTC TCTTCTACTT | 13980 |
| GAACTGCCTT TTGGAGGTCA CTGTTGTAGG CTGCAATGAT TTCAGCTGC AATTCAGCAT | 14040 |
| CCACGTGAAG CAATTCCTT TCTGCTTTT CTTTACCGAC AGCAGCAACG ATTTCTTCTT | 14100 |
| GGAAGGCAAT CAATTCCTT ACAGCTTCGT GCCCTTTAAG GAGCGCTTC AACATGATTT | 14160 |
| CTTCTGACAA TTCTTTGGCA CCAGACTCTA CCATGTTGAT AGCGTGCTTG GTTCCAGCTA | 14220 |
| CTGTCAATTC AAGAAGAGAT TGCTCTGCTT GTTCTTGACT TGGGTGATG ATGATTTGGC | 14280 |

948

| | |
|--|-------|
| CATCTACATA TCCCACCTTGT ACCCCAGCAA TTGGTCCGTC AAATGGAATA TCTGAAATAG | 14340 |
| ACAGTGCCAA AGATGAACCA AACATAGCAG CCATTGGTGC AGATGCATTT TCATCATAAG | 14400 |
| AAAGCACTGT ATTGATGACT TGGACTTCAT TACGGAAACC TTCCGCAAAC ATAGGACGAA | 14460 |
| TCGGACGGTC AATCAAACGC GCTGTCAAGG TCGCATCTGT TGAAGGACGT CCTTCACGTT | 14520 |
| TCATAAAGCC ACCAGGAAAC TTCCCAGCCG CATACATTTT TTCTTCGTAG TTGACTTGGA | 14580 |
| GTGGGAAGAA ATCCCCAGTT GCCATTTTCT TAGACATAAC GGCAGCAGTC AAGACAGTTG | 14640 |
| ACTCACCGTA ACGTACGACA ACAGATCCAT TTGCTTGCTT AGCAACCTGA CCAGTCTCTA | 14700 |
| CAATTAATC ACGACCCGCA AAAGTCGTTT GAAACACTTG TTTTGCCATT TTAATCCCTT | 14760 |
| TTGGATTGAT GAAATTATAC GCCTTGCCTA CAAAGATCAA GATACCAAGG ACGTCAAAAG | 14820 |
| CAAAGTAAAA ATAGGAACT GACGAAGTCT TCGATGAAGA CAAGACAGTT TATCTTTTTT | 14880 |
| ACACAGCTTT TCGGCCGTGT TCAATTACAC AAGATATTTT GGACGGTTCG GCTTGCCGAA | 14940 |
| CATTTCTGTA GAAAAATAGG AAGGTGACGT CGCACTCGAC GAGTGCTAGG AAGCTTATCT | 15000 |
| TTTTTCTTAA GAAATGAGAC CAAAATTCAA GTCATCAAGA TACCAAGCCG TCAAGCAACT | 15060 |
| CAAAGGAAGA TAGGAAATCG AACGACGGAG CGACTACTCC TAGGGAGATT TATCTTTTTT | 15120 |
| CACAGAGTTG TAGGCAAGTT CAGTTTTCAT GATACATCAT TAGAAAGGTT TAATACTAAA | 15180 |
| GTATCTAAAG CTTTCACGCT AATCGCTATC GGGCGATTAG CTAAATGCTT TACTAACTCT | 15240 |
| CTCGTCAAAT AACATCGATT TGAATCACTC GTGTCGTAA ATCTTACAGT TTAAATGCAT | 15300 |
| TGTATTATTT AATACCTTCA TCTTTGTATC AAGTACGTAC AGAATTTATT TTATCATATT | 15360 |
| TTTCTTAAAA AGTGAGGTCT TTACCATTAA AAAGGAACCA TTCCCCTCAC CTGAGAAGAA | 15420 |
| TGGTTTGCTT TTATTATCCT AGAGACTGGT GATTAAACAA GGCATGGGTT GCTTGATGGA | 15480 |
| TGTATTTTGC TGTATCAGCA TTATTCATCG TATAGAGATG CACACCGGCA ACATCCTGAG | 15540 |
| TTACCAAGTC CACGATTGG TCCACTGCAT AGGCAAGTCC TGCTGCTCTG AGCGACTCAG | 15600 |
| GGTCATGCTC ATACTTGTCT AAGATGGCTT TAAATTGCG TGGAAGATGG ATATTCTCAC | 15660 |
| AAGTCTTCAA GAGTCGGAGA GCCTGATTTC GATTGAGAA TGGCATAATT CCTGCATGAA | 15720 |
| TGGGAACATC AATCCCAGCC AAGATACACT TGTCTGAAA ATCATAGAAG CGCTCATTTG | 15780 |
| CAAAGAAGAG CTGAGTTACG AGGCTCGAAC AGCCTGCATC CACTTTCTTC TTAAGATTTT | 15840 |
| GAATATCTGA AATCTGATTT GCGGAATCTG GATGCCCTTC TGGATAGCAA GCTCCAATAA | 15900 |
| TATCAAAGTG AGGGGTTTGT TCCTTGATAA ACTCAATCAA GTCGGTTGCA TAGCGGAAAT | 15960 |
| CCTTTTGTGG TTCCACGTCT GGAATAATAT CCCCACGAAG AGCCAAGATT TTCTGCACCC | 16020 |
| CAACTTTGTC CAAGTCAGCA ATAGTTTCAG CAACCTTGTC CTTAGTTAGA TAAATAGCTG | 16080 |

949

| | |
|--|-------|
| GCAAGTGGGC AATGGTCGGA ATCGCCAAAT CATTTTGGAT AAAGTCAGCC AAACGAACCG | 16140 |
| TCGTTTCCTT GATATTAAAT TTATTATTGC TGGCAGTTAC ACTGATAAAA TGGGGAGCCA | 16200 |
| ACTCCTGCAT ATCCTGCAAG GCTGAAATAA TGTTATCATT ACCCACGGCT GGGTTTGGAG | 16260 |
| GGAACACTTC AAATGAGAGT GACGGTGTTC GCGTGACAT ATGTAATAAC CTTTCTAGT | 16320 |
| TGATTTCTTT TTGAACAACC ACTGTATGGA GAGAAATCCA ATCTTACAAT TTCTCACGCG | 16380 |
| CAGCTTTAGC TGCTTCAACA AGGCGGATCA AGCTTCTTTT TGTTTCTGGG ATACCACGTG | 16440 |
| TTTTCAAACC ACAGTCAGGG TTGATCCAAA CTTTCTTGCT TGGCACTTTA GCAAGGATGG | 16500 |
| CTTCGATTGT GTTGTGATTT TCGCCTTCAT TTGGTACACG AGGTGAGTGG ATATCGTAAA | 16560 |
| CCCCAGGTCC CACTTCTGTT TGGAAGTTTT TCGCTTTGAG TTCGTCCAAG ATTTCAAGGT | 16620 |
| TTGAACGGTT AGCTTCAAAG GAAATAACGT CTGCATCCAT GTTATCGATA GCTGGGATGA | 16680 |
| TATCTGTAAG TTCTGAGTAA CACATGTGAG TGTGGATTTG TGTGTCTGGC GCTACTGTTG | 16740 |
| AGTGTACCAA GCGGAAGGCA GGAATAGCCC AGTCAAGGTA GTCTTCGTAC CAGTCGCTAC | 16800 |
| GCGGGAGTGG CAATTTTTCA CGAAGAGCAG CCTCGTCGAT TTGGATGATT TTCACACCAG | 16860 |
| CAGCTTCAAG GTCAAGTACT TCATCCTTGA TAGCAAGGGC GATTTGGAGA GTTGAATCCT | 16920 |
| TGATAGAGAT GTCTTCACGT GGAATGACC AGTTAAGGAT GGTAACAGGT CCAGTCAACA | 16980 |
| TACCTTTAAC AGGTTTGTTC GTACGACTTT GTGCATAGCT AGACCATTTA ACAGTGATAG | 17040 |
| GGTTAAGACG AGTGACATCA CCCCAGATGA TTGGTGGTTT TACCCACGC ATACCGTATG | 17100 |
| ATTGTACCCA TCCATTTTGA GAGAAGAGGT ATCCTGACAA GTTTTGACCG AAGTACTCAA | 17160 |
| CCATGTCATT ACGCTCAAAT TCACCGTGAA CAAGGACATC AAAGTCAATA TCTTCTGCC | 17220 |
| ACTTGATCCA TTCGTCAATC GTTTCAGCAA GGAAAGCGTC GTACTCTTTT TGAGACAATT | 17280 |
| CACCTTTACG GTAAGCCAAA CGTTTGGCAC GAACTTCTTT TGTTTGAGGG AATGAACCAA | 17340 |
| TCGTTGTTGT TGGAAGAGCT GGAAGTTTGA AAGCTTCTTC TTGGATAGCT TCACGTTCTG | 17400 |
| CAAAGGCTGG CAAACGAGTG TAGTCTGCGT CTGTCAAGCC AGCGATACGC GCACGAAGTT | 17460 |
| CAGCATTTTC ACCAACACGC TCAGTCGCAA AGAGTTCTTT GTTGGCTGCA AGAGCTTCTG | 17520 |
| AACCTTGACC ATTTCTGGATA GCATCCAAGT CACGGATTTC ATCCAATTTT TCAACTGCAA | 17580 |
| AGGCAAAGTG GTTCAAGAGT GCTGGTTCAA ATTCTTCATT AGCAGTTGTA AATGGCACAT | 17640 |
| GAAGAAGTGA GCAAGAGCTT GTCAAAACGA TGTTTTCAGC TGGAATTTGC TCAAGAACAG | 17700 |
| CCAAGCTCTT TTCGTAGTTG TTGCGCCAGA TGTTTTCACC ATTGACAATA CCTACATAGA | 17760 |
| GAGTCTTGTC AGCTGGGAAG CCACCTTTAA CGAGTTCAAG AGTTTCTTTA CCTTCAACAA | 17820 |

950

| | |
|--|-------|
| AGTCAAGACC GATAGCATCT ACTGGTAAGT TTACAAGGTC AGCGTATACG TCACGAACAT | 17880 |
| CACCGAAATA AGTTTGAAGC AAGACTTCAA GACCTTTTTT GTCAGCCAAG AGTTTGTGT | 17940 |
| AAAGGTTCOA GAAGAGAGCT TTTTCTTCAG CTGTCAAGTC TTTTACAAGA GCCGCTTCAT | 18000 |
| CCAATTGGAT GCGAGTCGCA CCAAGTTCAG CCAATTTAGC AAAAAGTTCT TGGTAAGCAG | 18060 |
| CCACTAAGCT ATCTACGAAG TCGTCTGCTT TCACGCCTTC TTCAAAGTCT GACAATTGAA | 18120 |
| GGAAAGTGAA GGGACCTACA AGAACAGGAC GAGTGTTCOA TCCAAGTTCT TTGGCTTCTT | 18180 |
| GGAAGTCATC GAAAATCTTG TGACCAGCCA ATTTTACTTG AGTGTCTTTT TCAAAATTTAG | 18240 |
| GAACGATGTA GTGGTAGTTA GTGTTGAACC ATTTCTTCAT TGGAAGGGCG CGAACGTCCC | 18300 |
| CTTTTTCTCC CTGGTAACCA CGTCCCAAAG CGAAGTAGCG CTCAAGGTCA GACAAGTCCA | 18360 |
| AGTTTTGAAC GGATGCAGGC ACCACGTTGA AAAGGAAAGC CGCATCTAGG AAGTTATCAT | 18420 |
| AGTGAGAAAA GTCATTTGAT GGAATTCAG TGATGCCTTT TTCTTTGACA ATGTTCCAGT | 18480 |
| GTTTAGCACG CAAGTCTTTT GCTGCTGCTA AAAGTTCTTC TTCTGAGATT TCTTTCTAA | 18540 |
| AGTATTTTTT AGTTGTAAAT TTAAATTCAC GGAATTCGCC CAAACGAGGG AAACCGATGA | 18600 |
| TTGTAGTTGA CATGATGTGT CCTCCAAAAT TTGTTGTTGA AACTATCTTA ACAGAAAAGA | 18660 |
| AAGCGTCTGT ATAATTGTAA AAAATTAGGG TTTGATATAG TTTGAAACTA TATATCTGTT | 18720 |
| TCGGACAAAA GAAAAGACT TGAAGCAAAC GTCTCAAATC CTTTGTAAAT CTACTTTTAC | 18780 |
| AGCTATATTC CAATTAGAAT ACTAAACAT GTTATTAGTA ATTCTTATAA GTGACTATGA | 18840 |
| CCTGTATTAT GAAAAGACTA TAACTGATTC TAGTCAACTT TTTCCCTGTT CAAGTGGGAC | 18900 |
| GATTGCTAGT GTCTTTCCTA AACTGGCTAG GACTTTTAAAG ACTGTATCCA ACTGAGGACT | 18960 |
| AGTCTTTCCT GTCTCCATCC TAGCTATGAC AGGCTGGCTT ATCCCACTGA CTTCTTCCAG | 19020 |
| CTTTTCTGA CTGATTCCTT GTTCATACCT AGCCTCAATC AACTCGCTCA TGATAGCCAC | 19080 |
| TCGCATATCA CTTTCAAGGA TTTCTCCTT GCTAAAGAGC TCAGATGGAC ATCCTTCCAA | 19140 |
| TTACTCCCAA TAGCACTATT CTTCACTACT TAACCCTCTT TTTTPTACGT CTATGTATTT | 19200 |
| TTAAAAAAT GAGCGAATTA TGATTCGATA GATTGACCAG TGGGTTTAAA GTTGGTGCTA | 19260 |
| GCCTATTTCT TAAGCGATTT TCCTTTTCTA GGATAAAGCA GTTCCTGCTT GCTTAACCCC | 19320 |
| AATTTTCCAC GATGAATCCA ATAGTAAATG GTTGAAATTC CCACGTAAAC CCCTTTAGCC | 19380 |
| ATCACCATCA TTTCAGGCGA AAATTTTGG TTATGTTTTT GGTATGTAT AGTGGAGAAT | 19440 |
| CTTTTCTTTT AGTTTCTTAA GACTGTTGAG CGTAGTCGGC AGAATAAATC TCTTTGAAGC | 19500 |
| GCCCTTTTCC AAGACATTGT CGGACTGTCC CACGCTTGAT TTCAGTGTGG ATAGTTTGAG | 19560 |
| GAGCTTTTCC AAGTAGAGAG GCAATTTCTC TATTTGATTT TCCTTCTTTT TTCCATCGTT | 19620 |

951

| | |
|--|-------|
| CGATTAAGCG ACGGCTATCG ATTGTCAAAT GTTTCGCTTT TGTAGTATAA TTGCTTTGCA | 19680 |
| TTTCTGTGCC TTTTAATCAT TTCAATCTTA AATTGGACTT TTTTACTTG GGTGTACTT | 19740 |
| AATCTATGAG GAAGACAAGA AAAAGAATAT CAATCAAGTA AAGTCACAAA GTCACATTAG | 19800 |
| CTCCGAGCAA CCATTGCAAA TTGAGGTACT CACACAATGA TTAAAACATT TCTCTCTGCC | 19860 |
| CTTTCGGTCA TTCTCTTTTC TATCCCTATC ATAACCTATT CTTTTTTCCC ATCTTCTAAT | 19920 |
| CTTAACATTT GGCTATCTAC CCAACCTATC TTGGCAGAGA TTTATGCCTT CCCCTTAGCT | 19980 |
| ACTGCAACTA TGGCTGCTAT TTTAAGTTTC TTATTTTTTT TCCTATCTTT TTACAAGAAA | 20040 |
| AATAAACAAA TACGGTTTTA CTCTGGCATT TGCTCTTAC TATCGCTCAT ATTACTATTA | 20100 |
| TTCCGAACAG ATAAAACCTT TTCTTCTGCA TCAAATAAGA CTAAAACCTT AAAATTAGTA | 20160 |
| ACTTGGAACG TCGCTAATCA AATAGAAGCA CAACATATTG AGCGAATTTT TAGCCATTTT | 20220 |
| GACGCCGATA TGGCTATATT CCCTGAACTA GCTACCAATA TCAGAGGTGA GCAAGAAAAC | 20280 |
| CAGAGAATCA AACTATTGTT TCATCAAGTT GGACTTTCTA TGGCCAACTA TGATATTTTC | 20340 |
| ACTTCTCCAC CTACCAATAG TGAATAGCT CCTGTGACTG TGATTGTCAA GAAAAGTTAT | 20400 |
| GGTTTCTATA CAGAAGCTAA AACTTTTCAT ACAACACGGT TCGGGACAAT TGTATTACAT | 20460 |
| TCGAGAAAAC AAAATATACC AGATATCATT GCCTTGCATA CTGCGCCTCC TCTGCCAGGT | 20520 |
| TTAATGGAAA TCTGGAAGCA AGACTTAAAC ATCATTCATA ATCAATTGGC TTCAAATAT | 20580 |
| CCAAAGGCTA TTATTGCAGG TGATTTTAAT GCAACTATGC GTCATGGAGC ACTTGCAAAA | 20640 |
| ATAAGCTCTC ATAGGGACGC ATTAAATGCA CTGCCACCTT TTGAAAGAGG AACTTGGAAT | 20700 |
| AGCCAAAGTC CAAAACCTTT TAATGCAACA ATAGATCATA TTTTATTGCC TAAAAACCAC | 20760 |
| TACTATGTTA AAGATTTAGA CATTGTAAGT TTTCAAACT CTGATCATAG ATGTATTTTT | 20820 |
| ACAGAAATCA CATTTTAATT ATTTTATATA AAATCACCCC TCTAATGTTT ATAACTAGA | 20880 |
| GGGGGAATTT GTATCCTACT ATCGTTTAAAC GCACTTCTGC ATTGACTTTT TCTTCGAGAG | 20940 |
| ACGCTTGGAT TTTTTCATA TAGCGTGCAG CTCTTCGTC CGTTAAGCTG TCTTCTGGAT | 21000 |
| TTTGGAAGGT CAAGCTATAA GCCATTGACT TCATACCAAG TCCCAGTTTT TCACCTGAGA | 21060 |
| AGACGTCAA GAGTTTGATA TCTGTCAAAC GTTTCACGCC GGCAGCTTGG ATAGCATCTA | 21120 |
| CAACTTCTTG GTGAGTCACT TCTGCCTTGA GGAGAAGGGC AACGTCACGG CTGACTGCTG | 21180 |
| GGAATTTGGT GATTTCCACA AATGGAACAG CAGGTTGGAG CGCCCCTTCG ATGGCTGAAA | 21240 |
| GGTTAAGCTC AGCTACATAC GTTCTGGA TATCGTAAGC CTTGGCAGTG ACTGGATGCA | 21300 |
| CTTGGCCAAG GAAACCAAGA ACTTGGTCAC CGAGTGAAAT CACGGCTGTA CGACCTGGAT | 21360 |

952

| | |
|--|-------|
| GAAGGCTAAC GATTTCAGAT GTTGCTGTAT AGGTTACTTG GAGTCCCAA CGAGTAAATA | 21420 |
| GGGCTTCAAG GATTCCCTTA GCATAGAAGA AATCAACTGG AACTGCTGCT GTTTGGAAAT | 21480 |
| CTTTTTCAGC AACCAAGCCT GTCAAGGCAA AGGCAAAGCT GTTGATCTCA TTTGGAAGTT | 21540 |
| CTTCTTTTGG ATTACCTGTT TGTTCAAAGA CTTTTC CAAT CTCATAAAGG GCCAAGTTTT | 21600 |
| TATTCTTACG AGCCACGTTG TAGGCAACGG TATCAAGGAT CCCTGAAATC ATATTTTGAC | 21660 |
| GGAGGACTGA ACGATCCACA GTCATTGGCC ACATGAGTTC AGTAAGGTTA CTGGTTGAG | 21720 |
| CTGTGAACTC AACTGCTTTT TCAGGAGTTG TCAGAGCATA GGTGATGATT TCTGTCAAAC | 21780 |
| CTGCTCCTTC AGCAATGGTA CGAACTTGAC GCGGAGTTT TTGTATCACA GTCAATTACAC | 21840 |
| CAGCTGTACC ATCGTCTTTT GGAAGGCTGG TTGGCAAGCG GTCATATCCA TAGATACGAG | 21900 |
| CGATTTCCTC AAAGAGATCA GCTTCGATTG TGATATCCCA ACGACGACGT GGTACGCTGA | 21960 |
| CTGTAAAGCT GTCTGCATTT CCAGAAAGAC CAAAGCCAAG ACGACGGAAG ACGTCTTCTA | 22020 |
| CATCAGCATA AGACAGCTCA GTTCCGAGGA CACGGTTAAC ATCAGCAAGG GTTGAAGAAA | 22080 |
| CTTCCACATC AGAGGTATCA AGCTCACCCG CTGAAACGAT ACCCTTACGC ACCGTCGCGC | 22140 |
| CTGCAAGCTC TGCAATCATG CTAGCTGCCG CATCAAGGGC TTCATTAACT GTTGCCACAT | 22200 |
| TAATTCCTTT TTCAAAGCGA GAAGATGACT CAGAACGAAG GTTCAGGCGA CCACTTGTCT | 22260 |
| TACGGATAGA TTTGCCATTA AAAACAGCAG CTTCAAGGAT AACACGACTA GATTTTTCAG | 22320 |
| AAATTTCTGT AGCCTGACCA CCCATAACAC CGGCAAGGGC TACTGGTTTG TCAGCAACTG | 22380 |
| TAATCACGAG GTCTGTCTCA GCCAAGTCTC GTTCTTCACC GTCCAGGGTC ACTAATTTTT | 22440 |
| CACCATCACG CGCTTCACGC ACACGGATGT CAGTCCCTTC AAATGTGTCC AAGTCAAAAG | 22500 |
| CATGCATAGG TTGACCAAAG TAGAGCAGGA TGTAAGTTGT CACGTCTACA ACGTTATTGA | 22560 |
| TGGGACGGAT GCCTTCGTTT ATGAGAAGGT TTTGCAACCA TTGTGGACTT GGTGCGATAG | 22620 |
| TCACATTGTC CAAGATACGA GCTGCATAGT AAGGCGCCTT GTCTGTCTCA ATGCTGACAG | 22680 |
| AAAGGGCATC TGCCGCAGCT TCATTAGTTT CTGTTAGAGT AAATTTTTTA AAGTTGACTG | 22740 |
| CCTTGTCTA GATGGCTGCC ACTTCGTGAG CCACTCCACA CATAGAAAGG GCATCTGCAC | 22800 |
| GGTTTGGTGT GATGGAAAGT TCGATGATTT CATCATCCAA GTCTAGGTAA GAAAAGACTT | 22860 |
| CCTCACCTGG CACGGCATCT TCAGGCAAGA TTTGGATGCC ATCTGCGAAT TCCTTAGGCA | 22920 |
| CAACTGAGTC AGAAATTCCC AATTCACCAA GTGAACAGAT CATTCCAAGT GACTCCAAAC | 22980 |
| CACGGATTTT TCCTTTTTTG ATTTTGTAGT TATCAGCGAT ACGAGTCCT GGAAGAGCCA | 23040 |
| CCATGACCTT GATCCCAGCA CGCACATTTG GGGCACCACA AACGATCTGA CGCTCTTCTT | 23100 |
| CTTCGCCAAC GTTAATCTGA CAAACATGGA GGTGAGTCTC TGGCACATCT TCGCAAGACA | 23160 |

| | |
|--|-------|
| AGACCTCACC GACGACAATT TTTGAGAGAC CAGCAGCTGG TGATTCGACA CCCTCTACCT | 23220 |
| CGATCCCTGT AGTTGACATT TTTTCAGCCA ACTCTTGTA TGGCACATCA ATGTCCACCA | 23280 |
| ATTCTTTTAA CCATTTATAA GATACAAGCA TAATTTAGTT CTCCAGAATG ACAGTTGTCA | 23340 |
| CTCTAGTTCT TTTCCTTTCC TATCATTTCA ATAGAAGAAT CCTCTTCTTA CCTTAATTTT | 23400 |
| TTTCTCAGTA ACCAATCCGT ATCTACTTTT TGACCAACCA TAAAATGATG TTGGCTAAAT | 23460 |
| TTTTCAAAAC CATATCGGTT ATAAAACGCT TGAGCTTTTG TATTATGCTC CCAAACACCT | 23520 |
| AGCCAAGCCC AAGAAAACT ATTTTTTGTA GCAAGTTCAA GTGCGAATTC AAACAGTTGC | 23580 |
| TTACCTAGTC CAAATCCTTG GAATTTTGT AGCACATAGA GACGTTGAAT TTCAAAAGCG | 23640 |
| TCCTCTAATT CTCTCTCAGT TTGAGCACTT CCCAGTTGA CTTTGAGAAA ACCAGCTATC | 23700 |
| TCCTCCTCAT GCATAATGAA ATAGGTTTCA GAGTCAGGAT TTCCCAACTC AGTTGACAAA | 23760 |
| GTTTTCAGAC TATAAGCCTC TTCAAAGTAT TCCTGTAAC TCTCTTCCGT ATTATCATAC | 23820 |
| GCAAAGGTTT CACGAAAGGT TTGTTTGCA ATTTTAGCCA ACACCTCAAC ATCTGCCATT | 23880 |
| TCTACTTTTC TAATCATTAT TTAACTGTT CTGAGAAGCG GACATCTCCT TGGTAGAATC | 23940 |
| CACGGATATC GTTGATTCCA TAACGGAGCA TAGCTACACG CTCTTGTTCCA AGACCAAAGG | 24000 |
| CAAAGCCAGA GTATACAGTC GCATCGATAC CACTCATTTT AAGGACACGT GGGTGAACCA | 24060 |
| TACCGGCCCC CATAATTTTC ATCCAACCTG TTTTCTTACA TACATTACAG CCTTCTCCAC | 24120 |
| CACACTTGAA GCAAGAAACA TCCACCTCAA CAGATGGCTC TGTGAATGGG AAGTAAGATG | 24180 |
| GACGCAAACG AATTGACGC TCTTCACCAA ACATTTTTTG GACAATCAAC TGAAGCGTTC | 24240 |
| CTTGAAGATC AGCCATAGAG ATATTTTTC CAACTACCAA GCCTTCGATT TGGTGAAT | 24300 |
| GGTGACTGTG GGTGCGATCG TCCGTATCGC GACGGAAGAC ACGCCCTGGC GAGATCATCT | 24360 |
| TCAAAGGACC TTTAGAAAAA TCATGGGCAT CCATAGCACG CGCCTGAACT GGAGACGTGT | 24420 |
| GGGTACGGAG CAAGATTTCT TCAGTGATAT AGAAAGTATC CTGCATATCA CGAGCTGGGT | 24480 |
| GGTCTTTTGG AAGGTTTATA CGTTCAAAGT TATAGTAGTC TTGCTCCACT TCAAAACCAT | 24540 |
| CCACGACTTG ATAACCCATA CCGATGAAGA TATCTTCGAT TTCTTCACTG GTTTGTGTCA | 24600 |
| AAACGTGACG GTGACCAGTC GCAACTGGAC GACCTGGAAG CGTCACATCT ATACTCTCGC | 24660 |
| TAGCCAGTTG AGCCGCGACT TTCTTTTCTT CCAAGAGCTT AGCTGTTTCT TCAAAAGCAG | 24720 |
| CAGTCAAGAC ATCAGGAGCT TCATTGACGT GTTTCCCGAT GATTGGACGC ATCTCAGCAG | 24780 |
| AAACATCTTT CATCCCTTTG AGGATTTTCA TGAGCGAACC CTTTTTACCA AGGACAGAGA | 24840 |
| CACGCAAATC TTGCATCTCT TTTTCATTTT CAGCAGTAAT CTGCTTCAAG CTAGCCAGCG | 24900 |

954

| | | | | | | |
|-------------|------------|------------|------------|-------------|------------|-------|
| TTTCTTCGCG | AAGCGCTTTT | AATTGTTCTT | CAATAGTTGA | CATATTTCCCT | CCATCAGTCT | 24960 |
| CTCGTAGATA | AAAAGAAAAC | CACATGCCAA | AAACTCCACT | CGGAGCGTTG | ACACGCGGTA | 25020 |
| CCATCCGTTT | TCATCTGACA | AGTCAGACCT | TCATTTCTAA | ATCCATGCGC | AAGTGAATTC | 25080 |
| ACCCAGCTTT | CATATAGAGA | GCTTGCAGTC | ACGGCTCTCC | TCCCTGATAT | ACTTCCCTTG | 25140 |
| GGCTACTAGT | CTTTCAGATT | CCTATTCAAT | TACTACTTAG | TTTATCAGAT | TTTTACCATT | 25200 |
| CTTGCAAGAC | CTATCTTACT | TCTGCTTGT | AGCTTATTCT | TATCTAAATT | TATATAAACc | 25260 |
| TTATCTAAAT | TAACTATTTA | TAATTTTGT | AACAAAATTA | AATTAATTGA | CACTCCCCTA | 25320 |
| TAAATAAAG | AAGTTTAGAA | TTTAATGTCT | TCCAAACTTC | TTTATTCAT | ATTTAATGAA | 25380 |
| ATGCCACCTT | AACCGTGATA | ATAGCTAGTC | ATCAATAAAA | AACTATTTGA | ATAAGGATTC | 25440 |
| TCCATTGAT | TCAATCACCT | CTTATACCA | AGTAAAAGAC | ATTTTCTTAT | ATCGATTTAA | 25500 |
| TGTACCACTT | CCATCATCGT | TTGATCAAC | ATAAATGAGA | CCGTACCTTT | TAGAAAGTGT | 25560 |
| TGCAGTGGAC | ATAGAAACAC | AGTCAATACA | TCCCAAGAC | GTATAGCCCA | TAATTTCAAC | 25620 |
| ACCATCCTGT | AGAGCTTCAG | CAACTTGCAA | TAAATGTTCT | TTCATATACT | GAATTCTATA | 25680 |
| ATCATCTTGG | ACGGTTAAGT | TATTAAGTTC | ATCTTTTATT | AGTTGATCTT | TAGCACCTAA | 25740 |
| TCCATTTTCT | ACTATAAATA | ATGGGATTTG | ATAACGGTCA | TAATATCTAT | TTAAAATTAT | 25800 |
| ACGTAGTCCA | ATTGGATCAA | TTTGCCATCC | CCACTCTGAA | GACTCTAAAT | AAGGATTTAC | 25860 |
| TAAACCACCA | ATAATATTCC | CTTCTCCTGA | ATTATACTGT | GTTGGAAGAG | CAGATTGAGT | 25920 |
| CACACTCATG | TAATAGCTAA | AGGATAAAAA | ATCTACGGTA | TAATTTTTTA | ATAACTCTGC | 25980 |
| ATCTTCAGCT | GCAAACTCTA | TGTTAATGTC | ATTTTCCTTA | AAATATCTTT | TTGCATAATT | 26040 |
| CGGATAATAA | CCTCTAACAT | GCACATCTGA | AAATAGATAA | TTTAGATTCT | CATACTCATG | 26100 |
| AGTCGCCCCAT | ACATCTTTTG | GATTTGGAGT | CATTGGATAA | GCTGGCATAG | CTAATACCAT | 26160 |
| ACATCCCACC | TTAAACTCTG | AATTAATCTC | ACGAGCAATT | TTTGTAACCA | AACTTGAGGC | 26220 |
| GACTAATTCA | TGATGTATAG | CTTGATATAA | TTCTTGTTTC | GAAAGATTCT | CCTTAGGTAT | 26280 |
| ATCTATTCCT | CCACTAGTAA | ATGGTAATTC | CAAAACAGAG | TTTACTTCGT | TAAATGTAAG | 26340 |
| CCAATATTTA | ACTTTATCTT | TATACCTTTC | TAAAACGTT | CGAGCAAATT | TTTCATAAAA | 26400 |
| ATGAATCATT | CTCCTATCAA | CCCATCCATG | ATATTTTCTT | GCTAAATATA | ATGGAGTCTC | 26460 |
| ATAGTGTGAA | AGAGTTACAA | GTGGTTCTAT | CCCGTGAGCA | TGTAGTTCAT | CAAACAATTC | 26520 |
| ATCATAATAT | TTCAACCCAG | CTTCGTTAGG | TTCTTCCTCA | TCTCCTTTTG | GAAAAATTC | 26580 |
| ACTCCATGCA | ATAGAAGTAC | GAAAAACATT | AAAGCCCAT | TCAGAAAACA | AGGATATATC | 26640 |
| TTCCCTTATAT | TTATGATAAA | AATCAATACC | TATCAATTTT | AAGTTATCTT | CTGTAGGATT | 26700 |

955

| | |
|--|-------|
| TTCTGTTGCT TCTCCTAATC CACCTTTGGG TAACACATCC TGAAGTATA AGCCCTTACC | 26760 |
| ATCTTCATTA TATGCTCCCT CTACTTGATT AGCTGCAACA GCTCCACCCC AAAGAAAATC | 26820 |
| ATCTGGAAAA ATGGTCATAA CTTTCTCTCA TTATAATATT ACCAGTAATT CCTTAGAATG | 26880 |
| CTCGATTGTC TGATTATTAG GTAATACTAA TACATCTAGA AAATCATTGG TATTCCTTAC | 26940 |
| AATTACTGGT GTAAGTGTTC CGTAGCCTTT AGTCTTGATT AAATCAAGT CCATTTCAAA | 27000 |
| AATCAACTGA TTTTGGAAAA CTCTGTCTCC TTCTTCTACA TGAATAATA AACCTTGACC | 27060 |
| TTTTAGCTCA ACAGTATCTA ATCCAATATG AATTAGTAAC TCAACACCCT CATCACTCTT | 27120 |
| CAATCCAATT GCGTGCTTAG TCGGAAAAAT ATTTGTAATT TTCCCACAA ATGGTGCATA | 27180 |
| AACCTTACCT TCACTGGGA TAATCGCTAC TCCGTCTCCA ATTAGTTTAT CTGAAAATGT | 27240 |
| TTTATCCTGG ACATCGCTTA ACGGAATGAT TTCTCCTGAT ATAGGAGAAA ATATCATTTT | 27300 |
| TTTATTTGAA ACTCCAGCTT CAACTTCTAA ATTGCTAGAA CTCTCTTCTT CATCGATTCC | 27360 |
| AAATATATAA GCTAATACAA AGGTAATAAC AACCAGAAATG ACCGCCACAA TTAAAGCATT | 27420 |
| TACAATATTT GATGGCACAT CAGAATAAAT AAATTGAGGC AACGCTATCA AAGATGGGAC | 27480 |
| AGCAAATAGA TATGCTTTAA CACTAGTAAG ACCTGCAAAT AATCCCGCTA ATCCACCACC | 27540 |
| AATCATAGCT GCATAAAGCG GTTTTTTATA TTTTAAAGTC ACACCATATA ATGCAGGTTT | 27600 |
| GGTAATCCCT GCAAGTAAGG CTGAGAAACC TGCTGCAAAA GCAATTGTGT TTGTATTATT | 27660 |
| ATTTTTACTC TTTAATGCAA CAGCCATCGA AGCAGCCCTT TGAGCTAAGT TTGACCCTAA | 27720 |
| CATTGCTGGA AGAATTAATA CGTCTGGAGT AGCAATAGAT GCCGCCAAAA AAATAGGTGC | 27780 |
| AAAAGCCCAA TGCATTCCAG TCATAACAAT AAATGGCATA ATAGCACCAA GAATAGCTAA | 27840 |
| TGTAAGCCAT CCAGCTACAC CATACATTTG CCCAACTAGA TTTGATAATC CTTACCAAC | 27900 |
| AATTACTCCA ATAGGTCCGA CTACAATAA GGCAATACAG CTTGATACTA ATAATACTAG | 27960 |
| CGTAGGTTGC AAAAACTCT TAGTAATAGC TAGTGTTAAT TTAGCAATTA TTTTTCAT | 28020 |
| ATATTTTCATC AACCAAACCA TAATAAGAAT TGGAACGACT GATGAACCAT AACTAGCTGG | 28080 |
| TGTCACAGGT GCACCAAATA AACTAAGAGG ATTCCCTGAT TGCACCATTG GAACAAAATT | 28140 |
| TGGATGGAGA AGTACACCTG CTACAGACAT AGCTAATGTA GATGTTACTT TTAATTTTGT | 28200 |
| TGATGCAGAA TAAGCTAATA ACAGCGGTAA GAAATAATAT GGAGCATCCC CAAAAATGT | 28260 |
| CAAAAAAGCA ATAGTCTGAG AATCTGATTG CAATATACCA AGCATTTGTA AATGATTAC | 28320 |
| CAAGACTTTC AACATACCTC CCCCTAACAT TGCTGGAATG ATTGGAGTCA TGGAACCCAGC | 28380 |
| GATATACTCA ATGATTCTTT CTAAAATATT CCCTTTGTGC CCTTGAACAA CTGAATCGGA | 28440 |

956

| | |
|---|-------|
| TTCAAAATTG CCAAGTTTAA CGAATTCCTT ATAATAATTA GCTACATCAT TACCAAGTAT | 28500 |
| AATTTGATAT TGTCCATTCT TTTTCATAAT ACCTATTACA CCTGGTATCT TCTTCACATC | 28560 |
| ATCATCATTG ACTAAATTTT CATCTTTTAA TTCTAATCTT AAACGTGTTA CACAATGGGT | 28620 |
| AACTCTATTG ACATTTTTTT CACCTCCAAT TACATCGAGG ATTTTTTGTA CCGTATCTTT | 28680 |
| ATAACTCATG GTATTCTCCT ATTCTATTAA TCTAAATTTT TTGTTAAGCG ACGAATATGA | 28740 |
| GCCATCAAAT AACTAATTC ACTAGAAGTC AGCAAATAAT TGTACTCCGT TTGTATAAAC | 28800 |
| ATTGCTACCT GTTCACCACA TTCATATTCT CTAGGATATT TATTTTTCAT TAATGCTAAC | 28860 |
| AAGTCTTCAT CATCATCGTC GG | 28882 |

(2) INFORMATION FOR SEQ ID NO: 141:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 12835 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 141:

| | |
|---|-----|
| GCCTATGTCT TTTTCAAAAA AATGCTTGAC TTGAGACGGG AACTAGGGAA GTCTAAAGGC | 60 |
| GGAAGGCATT GATTTATACT CTTGCAAAAT CTCTTCAAAC CACGTCAACG TCGCCTTGGA | 120 |
| TTATATATGT AACTGACTTC GTCGATGCTT ATCTACAACC TCAAAGCAGT GCTTTGAGCA | 180 |
| ACTTGCGGCT AGTTTCCTAG TTTGCTCTTT GATTTTCATT GAGTATTATA TTACTTTCTA | 240 |
| TTTGTAGGAG GTGGCTTATG AAGATTCCTC TCTTAACTTT TGCAAGGCAT AAATTTGTTT | 300 |
| ATGTCTTGCT TACTTTGCTT TTTCTTGCTT TGGTTTATCG TGATGTTTG ATGACTTATT | 360 |
| TCTTTTTTGA TATTCATGCG CCCGATCTAG CTAAATTCGA TGGACAAGCA ATTAAAAATG | 420 |
| ACTTATTAAA ATCAGCATTA GATTTTCGTA TTCTCCAGTT CAATCTAGGT TTTTATCAAT | 480 |
| CATTATATAT TCCAATCATC ATTGTTTTGC TAGGTTTTCA ATATATTGAG CTGAAAAATA | 540 |
| AAGTTTTACG ATTGAGTATT GGAAGAGAAG TGAGTTATCA AGGGTTAAAA AGAAAGTTGA | 600 |
| CTTGCAAGT TGCAAGTATC CCTGTGTTGA TATATTTAGT GACTGTGCTG ATAATTGCAA | 660 |
| TTATAACCTA TTTCTTTGGG ACTTTTTCTC CTCTTGATG GAATTCCTA TTTCTGATG | 720 |
| GAAGTGGTTT ACAAAGACTC CTAGATGGAG AGATAAAAAG CTATTTGTTC TTTACTTGTC | 780 |
| TCCTACTAAT CGGTATTTTC ATCAATGCAA TCTATTTTTT ACAAATAGTT GATTATGTGG | 840 |
| GGAATGTGAC TCGTTCGGCA ATCACCTATT TGATGTTTCT TTGGCTTGGT TCTATGCTGC | 900 |
| TTTATAGTGC CTGCTTAC TATATGGTTC CTATGACGAG TTTGATGCAA GCTAGCTATG | 960 |

957

| | |
|--|------|
| GGGATGTAAG TTTGATGAAA CTCTTTACTC CTTATATCCT TTATATTGTC CCTTACATGG | 1020 |
| TGCTTGAAAA ATATGAAGAT AATGTTTAAAG AATTTTAAACA ATATTTTGCT AAATAGAAAAG | 1080 |
| ATTGTTTTTAC TACTTCGTAT AGTTCTGATG ATGATTTTGA TAAACCATCT ATTGTCAACA | 1140 |
| GCGGTTCAAA AGCAGGATGC TGTATCTTT TTCAAGAGAG AATTGATTTC AATTTTTTCC | 1200 |
| TATAATGACT ATTCTGAAGC GAATTTAGAA ATCCCCAAAC TATTGTTAAA CCTTTCGCTT | 1260 |
| TTTCATGGTAG GATGGCTCTC TGTCATTTTA CTGAAAGTG ATTTGGCAGA CCATTACCAT | 1320 |
| CACCTGATTC GCTATCAATC AAGCTCCTTT TTCGATTATA CAAGGAAACG ATTGGTTGTC | 1380 |
| ATTTCTAAAT TTTTACTCA AGATTTGTTT GTCTGGTTTC TTGGTTTACT TCCTCTAGGA | 1440 |
| ATTCATTTCA AAACAGTCGC ACTTTTCTTT TTAATTGCTC AGTTAATGAT GTTGTACTTA | 1500 |
| CTACTGTCTT ATCTGATAGC ACTGATTAGT GCGGGCGCTG GTTTTTCCTT TTTTCTCTAT | 1560 |
| TTTTTAGCAT TTGTGGGACA AGAATGGATG ATGGATCATA TTGTAACAGT GTATTAGTA | 1620 |
| CTCTTAAGTT TATTAGTTAT GTTGATTGTT AGTCGCTTGG AAGAGAAATT TAAGAAAGGA | 1680 |
| TAAACGATGA GACTTGAAAT TATAAATGGA CAGAAAATTT ATGGGAAAAG ACCTATTTTA | 1740 |
| AATCAGTTGA ATTTGGTGT TCAATCAGGA AAAATTTATG GACTTAAAGG TGATAATGGA | 1800 |
| TCTGGCAAGA CGGTTCCTTT AAAGATACTT GCTGGTTATA TTAAGCTTGA CAAAGGAAAA | 1860 |
| GTCTTCAAG ATGGTAAAGT TTACGGGGTA AAAAATCATT ATATTCAGGA TGCAGGAATT | 1920 |
| TTAATTGAAA AAGTCGAGTT TTTATCTCAT TTATCCCTGA GAGAAAATTT GGAAGTGTTA | 1980 |
| AGGTATTTTT CATCTAAAGT TACGGAAAAA AGAATTGCCT ATTGGATTCA ATACTATGAT | 2040 |
| TTACAGGAAT TTGAAGACAT TGAATACCGT CATTTATCCT TAGGAACAAA GCAAAAAATG | 2100 |
| GCCTTGATTC AAGCCTTTAT TTCCTCTCCT TCTATACTCT TTCTCGATGA ACCTATGAAT | 2160 |
| GCTTTGGATG AGAAGAGTGT GAGGTTAACC AAACAGGTCA TTTTATCTTA CTGAAAAAA | 2220 |
| GAAAAATGGTC TGGTTATCCT GACGTCGCAC ATATCGGAAG ATATTCAGGA CCTTTGTACA | 2280 |
| GATGTATTAG TTGTCGAAAA TGGACATATA CAAATGTAAA GGATATACAA TCCTAGGAGA | 2340 |
| TGGCTTATGG CACATCTAAA ATCATTATT ACACGATATT CCAAGGTTTA TATTGGTTTA | 2400 |
| GTTCTGCTGA TCTGGCTGTC TTTCTTCTTT ATCCCTTGGG ATAAACCACT TCTGGGGATA | 2460 |
| AGGATTGACA TCTTCATCAT ACAGAAAATC TTGCTAGCTT TTGGAATTCT GTCCATTCTC | 2520 |
| ATGGCCTTGC TGTCCAAGAA AGTCAGTCTC TTTGTMTTTC GACTGATTTC CTGTCTTTCT | 2580 |
| CTTTGGATTA ACTTATTTAT CACATTGACC ATTTTGCCGA TTTTGGCAA TTAAACAGTC | 2640 |
| ATAAAAGTCG GAGAGGTTAG CTTGAAAACT AACCTCTTTT TCCTTTTCAA AATGGGGATT | 2700 |

958

| | |
|--|------|
| CTTCCTTGAA AATAATCAGT AATTGTGCTA AAATTAAAGG AACATTCTAA AATATTCGGA | 2760 |
| ATTTAAAGTA AGGAAAAACA TGGCTAATAT TTAAAAACA ATTATCGAAA ATGATAAAGG | 2820 |
| AGAAATCCGT CGTCTGGAAA AGATGGCTGA CAAGGTTTTC AAATACGAAG ACCAAATGGC | 2880 |
| TGCTTTGACT GACGACCAAC TAAAAGCAAA AACAGTTGAA TTTAAGGAAC GTTATCAAAA | 2940 |
| TGGAGAATCA CTGGATTCAT TGCTTTACGA AGCATTTCGCG GTTGTCCGTG AAGGTGCCAA | 3000 |
| ACGTGTCCTA GGTCTCTTCC CTTATAAGGT TCAGGTCATG GGGGGGATTG TTCTTCACCA | 3060 |
| TGGTGACGTG CCAGAGATGC GTACAGGGGA AGGGAACC TTGACTGCGA CCATGCCGGT | 3120 |
| ATACCTCAAT GCCCTTTCAG GTAAAGGGGT TCACGTAGTT ACGGTTAATG AATACCTGTC | 3180 |
| AGAACGTGAC GCGACTGAGA TGGGTGAATT GTACTCTTGG CTTGGTTTGT CAGTAGGGAT | 3240 |
| TAACTTGGCT ACCAAATCTC CAATGGAGAA AAAAGAAGCC TATGAGTGTG ATATTACTTA | 3300 |
| CTCAACTAAC TCAGAAATCG GATTTGACTA CCTTCGTGAC AACATGGTCG TTCGCGCCGA | 3360 |
| AAACATGGTA CAACGTCCGC TTAATATGC CTTGGTCGAT GAGGTTGACT CTATCTTGAT | 3420 |
| TGACGAGGCT CGTACACCTT TGATTGTATC AGGTGCCAAT GCGGTTGAAA CCAGTCAGTT | 3480 |
| GTATCACATG GCAGACCACT ATGTAAATC TTTGAACAAA GATGACTACA TCATCGATGT | 3540 |
| GCAGTCTAAG ACTATTGGTT TGTCTGATTC AGGGATTGAC AGGGCTGAAA GCTACTTCAA | 3600 |
| ACTTGAAAAC CTCTATGACA TCGAAAACGT GGCTTTGACT CACTTTATCG ATAACGCCCT | 3660 |
| TCGTGCCAAC TACATCATGC TTCTCGATAT TGAATATGTG GTGAGCGAAG AGCAAGAAAT | 3720 |
| CTTGATTGTC GACCAATTTA CAGGTCGTAC CATGGAAGGT CGTCGTTATT CTGATGGATT | 3780 |
| GCACCAAGCT ATTGAAGCCA AAGAAGGTGT GCCAATCCAG GATGAAACCA AGACATCTGC | 3840 |
| CTCAATCAGC TACCAAAACC TCTTCCGTAT GTACAAGAAA TTGTCTGGTA TGACGGGTAC | 3900 |
| AGGTAAGACT GAGGAAGAAG AATTCGGTGA AATCTACAAC ATTCGTGTTA TTCCAATCCC | 3960 |
| AACAAACCGT CCTGTTCAAC GTATTGACCA CTCAGACCTT CTTTATGCAA GTATCGAATC | 4020 |
| TAAGTTTTAA GCGGTGTGCG AAGACGTTAA GGCTCGTTAC CAAAAGGGTC AACCTGTCTT | 4080 |
| GGTTGGTACA GTAGCGGTTG AAAC TAGTGA CTACATTTCT AAGAAATTGG TTGCAGCTGG | 4140 |
| TGTTCTCTAC GAAGTCTTGA ATGCCAAAAA CCACTATAGA GAAGCCCAA TCATCATGAA | 4200 |
| TGCTGGTCAA CGTGGTGCCG TTACCATCGC AACCAACATG GCGGGTCGTG GTACCGACAT | 4260 |
| CAAGCTTGGT GAAGGTGTTT GTGAACCTGG AGGACTTTGT GTTATTGGTA CAGAACGTCA | 4320 |
| TGAAAGTCGT CGTATCGATA ACCAGCTTCG TGGACGTTCA GGTGTCGTAAG GAGATCCAGG | 4380 |
| TGAGTCACAA TTCTACCTAT CTCTGAAGA TGATTGATG AAACGTTTGG GTTCTGAACG | 4440 |
| CTTGAAGGGA ATCTTTGAAC GCTTGAACAT GTCTGAAGAG GCCATTGAGT CTCGCATGTT | 4500 |

959

| | |
|--|------|
| GACGCGTCAG GTTGAAGCAG CTCAGAAACG TGTCGAAGGA AATAACTACG ATACCCGTAA | 4560 |
| ACAAGTCCTT CAATACGATG ATGTCATGCG TGAACAACGT GAGATTATCT ATGCTCAACG | 4620 |
| TTACGATGTC ATCACTGCAG ATCGTGA CTT GGCACCTGAA ATTCACTCTA TGATCAAACG | 4680 |
| CACGATTGAA CGTGTCTTG ATGGTCATGC GCGTGCCAAA CAAGATGAAA AACTAGAGGC | 4740 |
| AATTTTGAAC TTGCTAAGT ACAACTTGCT TCCTGAAGAT TCTATTACGA TGAAGACTT | 4800 |
| GTCAGGCTTG TCTGATAAGG CCATCAAGGA AGAGCTTTTC CAACGTTCTT TGAAGGTTTA | 4860 |
| CGATAGTCAG GTTTCAAAAC TACGCGATGA AGAAGCAGTT AAAGAATTCC AAAAAAGTTT | 4920 |
| GATTCTACGA GTGGTGGATA ACAAGTGGAC AGATCATATC GATGCCCTTG ATCAATTGCG | 4980 |
| TAACGCGGTT GGACTTCGTG GCTATGCTCA GAACAACCCT GTTGTGAGT ATCAGGCAGA | 5040 |
| AGGTTTCCGT ATGTTTAATG ATATGATTGG TTCGATTGAG TTTGATGTGA CACGCTTGAT | 5100 |
| GATGAAAGCA CAAATTCATG AACAAGAAAG ACCACAGGCA GAACGTCATA TCAGTACAAC | 5160 |
| AGCGACTCGC AATATCGCTG CTCACCAAGC AAGTATGCCA GAAGATTGG ATTTGAGCCA | 5220 |
| GATTGGACGC AATGAACCTT GCCCATGTGG TTCTGGTAAG AAATTTAAAA ACTGTCACGG | 5280 |
| TAAAAGACAA TAAAATGAGA TAGTTTAGAG GCGGATATCT TGTGAAAAGT AAATTTTAC | 5340 |
| TGGGTATCCG TTTGCTTTAT AAGGAGATGA GTTATGGTAT TTACAGCAAA AAGCTCTAAA | 5400 |
| ATAAATATAG AAGAAGTTCG TGCCTTGTC AATTAGAAG GTCAGGCTTT GGAGAGGAAA | 5460 |
| TCACAGCGAG ATCAAGAGCT AGAAGCCATT ATACGTGGAG AAGACCAGCG AATCTCTTG | 5520 |
| GTAATCGGGC CATGCTCATC TGACAACGAA GAAGCTGTCC TTGAATACGC TAAGCGTTTG | 5580 |
| GCAGTCCTAC AAGAAGAAGT GGCAGATCGT ATCTTTATGG TTATGCGTGT TTATACTGCC | 5640 |
| AAACCCCGTA CCAACGGAGA TGGCTATAAG GGCTTGATTC ACCAGCCTAA CGCGACAGAA | 5700 |
| GCGCCTAGTC TTATCAATGG AATCAAAGCC GTTCGCCATC TTCACTATCG TGTATCACA | 5760 |
| GAAACAGGGA TGACAACTGC TGATGAAATG CTTTATCCTG AAAACCTTCC GCTTGATAGT | 5820 |
| GATTTGATTT CTTACATGGC AGTTGGTGCC CGTTCAGTTG AAGACCAGCA ACACCGCTTT | 5880 |
| GTGGCAAGTG GGGCAGGATT TTCTACTGGT TTTAAAAATC CAACCTCTGG AATCTCAAT | 5940 |
| GTCATGTTTA ATGGGATTTA TGCTGCTCAA AACAAACAAA GTTTCCTTTT CTTAGGAAAA | 6000 |
| GAAGTAGAAA CAACTGGGAA CCCGCTTCA CACGCTATTC TTCGTGGTGC TCTTAATGAG | 6060 |
| TATGGAAAAA ATATTCCTCA CTAATATTAT GACAATTTAA TTGATACCAT TGCCCACTAT | 6120 |
| GAGAAAAATG GCTTGGAAAA TCCTTTTATC ATCATTGATA CCAATCATGA CAATCTGGT | 6180 |
| AAGCAGTATA TTGAACAGAT CCGAATTGTC CGCCAGACCT TGATTAAACG TGCTTGAAT | 6240 |

960

| | |
|--|------|
| GAAAAAATTA AGCAGTTCCT TCGTGGTTTT ATGATTGAGT CTTATCTGGA AGATGGTCCA | 6300 |
| CAAAATGAGC CAGAAGTATT TGGTAAGTCT ATCACAGACC CTTGCCTGGG TTGGGATAAC | 6360 |
| ACAGAAGCTC TTGTCAGAGA AATTTACAAA ACGTTAGGAG AATAAGATGG CATTATTGTA | 6420 |
| AAAAGGTCAA GAAATCGATA TGGAAATCAT CAAGGCTGAA ACCCAATTGT CTGCGGAAGC | 6480 |
| CTTGAGACTC AAGGAAAGCC GTGACAGGGA ATTGGCAGAT ATTATTTCAG GGAAGATGA | 6540 |
| CCGTATTCTC TTGGTGATTG GTCCTTGCTC TTCTGATAAT GAAGAGGCGG TCTTGGAATA | 6600 |
| TGCTCGCCGT TTATCTGCCT TGCAAAAGAA GGTAGCGGAT AAGATTTTCA TGGTCATGCG | 6660 |
| CGTGTATACT GCTAAGCCTC GTACCAATGG AGACGGCTAT AAAGGATTAG TTCACCAGCC | 6720 |
| AGATACTTCT AAGGCTCCAA GCCTGATTAA TGGCTTGACG GCTGTGCGCC AGTTGCACTA | 6780 |
| CCGCGTGATT ACAGAGACTG GTTTGACAAC GGCAGATGAG ATGCTTTATC CGTCAAATCT | 6840 |
| GATCTTGGTG GATGACTTGG TCAGCTACCA TGCCGTTGGA GCTCGTTCTG TGAAGACCA | 6900 |
| AGAGCACCGC TTTGTGGCTT CTGGGATTGA TGCACCAGTA GGGATGAAAA ATCCAACCTC | 6960 |
| AGGAAATTTG GGTGTTATGT TTAACGCCAT CTATGCTGCT CAAAACAAGC AAACCTTCCT | 7020 |
| TTATCATGGG CAGGAAGTTG AGACATCAGG TAATCCTTTG GCCCATGTTA TCCTCCGTGG | 7080 |
| AGCAGTCAAC GAGTATGGCA ATTATATGCC GAATTACTAC TATGAAAATC TACTCCAAGC | 7140 |
| CATTGAACGC TATGAAACCA TGGGACTTGA AAATCCTTTT ATCCTCATTG ACACCAACCA | 7200 |
| TGATAACTCA GGCAAGCAAT ATATGGAGCA GATTGCAATT GTTCGCCAGA CCTTGCAGAA | 7260 |
| TCGTGATTGG AATGAGAAAA TTA AAAAGAC GGTTCGAGGA TTTATGATTG AATCTTACCT | 7320 |
| AGCAGATGGT CGTCAAAACC AACCAGAGAT CTTTGGTTGC TCTATTACTG ACCCTTGCCT | 7380 |
| AGGTTGGGAA AATACAGAGG CCTTGGTAGA AGAGATTAT GTTACCTTGA CAAAATAAGT | 7440 |
| GAAAAGGATG GAGTTGGGGA ATCTCAACTC CTTTGTATGA GAATGATAGT TGGACACGGA | 7500 |
| ATTGACATCG AAGAATTGGC TTCGATAGAA AGCGCAGTTA CACGACATGA AGGATTTGCT | 7560 |
| AAGCGTGAC TGACCGCTCA GGAAATGGAG CGCTTCACCA GTCTCAAAGC ACGCAGGCAA | 7620 |
| ATAGAATATT TAGCTGGTCG CTGGTCGGCT AAGGAGGCCT TTTCCAAGGC TATGGGAACG | 7680 |
| GGCATTAGCA AGCTCGGTTT TCAGGATTTG GAAGTCTTGA ACAATGAACG TGGGGCGCCT | 7740 |
| TATTTTAGTC AGGCACCATT TTCAGGAAAG ATTTGGCTGT CTATCAGCCA CACCGATCAG | 7800 |
| TTTGTGACAG CCAGTGTCAT TTTGGAGGAA AATCATGAAA GCTAGTCCAC ATAGACCAAC | 7860 |
| CAAGGCTCTG ATTCATCTGG GAGCTATTTCG ACAAATATT CAGCAAATGG GGGCTCATAT | 7920 |
| CCCTCAAGGA ACGCTCAAGT TGGCTGTGGT TAAGGCCAAT GCTTATGGTC ATGGAGCTGT | 7980 |
| TGCCGTTGCC AAGGCAATTC AAGATGATGT TGATGGCTTT TGCCTTTCCA ATATCGATGA | 8040 |

961

| | |
|--|------|
| AGCCATTGAA CTCAGACAAG CTGGACTCAG CAAGCCAATC CTCATTTTAG GAGTTTCTGA | 8100 |
| AATCGAAGCT GTTGCTCTAG CTAAGAATA TGACTTCACC TTGACAGTGG CTGGACTGGA | 8160 |
| GTGGATTCAA GCACTCTTAG ATAAGGAAGT GGACCTAACT GGATTGACAG TCCACCTCAA | 8220 |
| GATTGATTCA GGGATGGGAC GGATTGGTTT TAGAGAGGCA AGTGAGGTTG AGCAGGCTCA | 8280 |
| AGATTTGCTC CAACAACACG GTGTTTGTGT TGAAGGAATC TTTACCCACT TTGCTACTGC | 8340 |
| TGATGAGGAA TCAGATGACT ATTTTAATGC CCAGTTAGAA CGGTTTAAAA CTATTTTAGC | 8400 |
| TAGTATGAAG GAAGTTCCAG AGCTGGTTCA TGCTAGCAAT TCTGCAACGA CTCTTTGGCA | 8460 |
| TGTAGAGACT ATTTTCAATG CGGTCGTAT GGGAGATGCC ATGTATGGCC TCAATCCAAG | 8520 |
| TGGAGCGGTC TTGGATTGTC CTTATGATTT GATACCGGCC TTGACCTTGG AGTCTGCTCT | 8580 |
| GGTTCATGTC AAGACAGTTC CAGCTGGAGC TTGCATGGGC TATGGAGCAA CTTATCAAGC | 8640 |
| GGATAGCGAG CAAGTCATCG CGACCGTGCC AATCGGGTAT GCAGATGGAT GGACAAGAGA | 8700 |
| CATGCAAAAT TTCTCTGTCT TGGTAGATGG CCAAGCTTGC CCAATTGTCG GCAGGGTTTC | 8760 |
| GATGGACCAA ATCACTATTG GATTGCCTAA GCTTTATCCG CTAGGAACCA AGGTAACCTT | 8820 |
| GATTGGCTCC AATGGGGATA AGGAAATCAC TGCAACTCAG GTAGCGACCT ACCGCGTAAC | 8880 |
| CATTAATAT GAGGTGGTTT GCCTCCTCAG CGACCGTATT CCGAGAGAAT ATTATTAGAA | 8940 |
| AAGAAAGGAG TGGAGCATGA ATCTACATCA ACCCTTGCAT GTCITGCCCTG GTGTGGGACC | 9000 |
| AAAGTCAGCA GAAAAATACG CCAAAGTAGG AATTGAAAAC TTGCAAGATC TCTTGCTCTA | 9060 |
| CTTTCCTTTC CGTTATGAAG ACTTCAAAAC CAAGCAGGTG CTGGAGCTGG AAGACGGTGA | 9120 |
| GAAGGCAGTT CTTTCTGGTC AGGTAGTGAC TCCTGCTAGT GTCCAGTATT ATGGTTTCAA | 9180 |
| GCGCAATCGC CTGCGTTTTA GTCTCAAGCA GGGAGAGGTC GTTTTTGCGG TGAATTTCTT | 9240 |
| TAACCAGCCC TATCTGGCTG ATAAATAGA GTTGGGAGCA ACCCTTGCTG TCTTTGGAAA | 9300 |
| ATGGGACCGC GCTAAGGCTA GTCTGACTGG GATGAAGGTT CTGGCTCAGG TAGAAGATGA | 9360 |
| CCTCCAGCCT GTCTATCGTC TGGCTCAGG AATCAGTCAG GCCAGTCTGG TCAAGGTCAT | 9420 |
| CAAGACGGCT TTTGATCAGG GACTGGACCT CTTGATAGAA GAAAATCTGC CCCAGTCTTT | 9480 |
| ACTAGACAAA TACAACTCA TGTCCCGTTG TCAGGCAGTC CGTGCTATGC ATTTTCCAAA | 9540 |
| GTATTTGGCA GAATACAAGC AGGCTCTTCG CCGTATAAAG TTTGAGGAAC TCTTTATTTT | 9600 |
| CCAAATGCAG CTGCAGATGC TCAAGTCTGA AAATAGAGTT CAGGGAAGTG GTCTGGTTCT | 9660 |
| GAATTGGTCT CAGGAAAAAG TGACAGCAGT TAAAGTAAGT CTTCTTTTGG CCCTGACCCA | 9720 |
| AGCTCAGGAA AAGAGTTTGC AGGAAATTTT AACTGATATG AAGTCCGACC ACCACATGAA | 9780 |

962

| | |
|---|-------|
| TCGTCTCCTA CAAGGGGATG TGGGGAGTGG AAAAACGGTA GTCGCTGGCT TGGCCATGTT | 9840 |
| TGCGGCAGTG ACAGCAGGTT ATCAGGCTGC CCTAATGGTA CCAACAGAAA TCCTCGCAGA | 9900 |
| GCAACACTTT GAGAGTTTAC AGAACCTTTT TCCCAATTTG AAAGTGGCTC TCTTGACAGG | 9960 |
| TTCTTTGAAA GCTGCAGAAA AGAGAGAAAT CTTGGAGACC ATTGCCAAGG GTGAGGCTGA | 10020 |
| TTTGATTATA GGAACACACG CTCTGATACA AGATGGGGTG GAGTATGCTC GTCTTGGTTT | 10080 |
| GATTATTATC GATGAGCAGC ACCGTTTGG TGTAGGGCAA AGGCGTATTT TACGGGAAAA | 10140 |
| AGGTGACAAT CCAGATGTCC TCATGATGAC GCGACTCCC ATTCCACGGA CGCTTGCCAT | 10200 |
| CACAGCCTTT GGAGATATGG ATGTTTCCAT TATCGACCAG ATGCCAGCAG GTCGGAAGCC | 10260 |
| TATTGTGACG CGCTGGATCA AACATGAGCA ACTACCTCAG GTCTTGACTT GGTAGAGGG | 10320 |
| GGAAATTCAA AAAGGTTCCC AAGTCTATGT CATCTCTCCT TTGATTGAAG AATCAGAAGC | 10380 |
| TCTAGATTTG AAAAATGCCA TTGCCTTATC AGAGGAGTTG ACGACTCATT TTGCAGGCAA | 10440 |
| GGCAGAGGTG GCTCTTCTAC ATGGTAGGAT GAAGAGTGAC GAAAAAGACC AGATCATGCA | 10500 |
| GGATTTCAAG GAGAGAAAGA CGGATATTCT GGTTCGACG ACGGTTATTG AGGTTGGGGT | 10560 |
| CAACGTTCCC AATGCGACTG TCATGATTAT CATGGATGCC GATCGCTTCG GTCTCAGTCA | 10620 |
| ACTTCACCAG CTTAGAGGTC GTGTCGGTCG GGGGGACAAG CAGTCTACG CTGTTCTCGT | 10680 |
| TGCTAATCCC AAGACGGATT CTGGGAAAGA CCGCATGCGT ATCATGACAG AAACGACCAA | 10740 |
| TGGATTMTGC CTTGCGGAGG AAGATTTGAA AATGCGTGGT TCTGGTGAGA TTTTGGAAAC | 10800 |
| CAGACAGTCA GGAATTCCAG AGTTCCAAGT GGCTGATATT ATCGAAGATT TTCCGATTTT | 10860 |
| AGAAGAAGCA AGAAAGGTTG CTAGCTACAT TAGTTCTATA GAAGCTTGGC AAGAAGATCC | 10920 |
| AGAGTGGCGC ATGATTGCCC TTCATCTGGA AAAGAAAGAA CATCTGGATT AAGCTTTCTC | 10980 |
| TAAGGAAAAC TTATACTCAA TGAAAATCAA AGAGCAAACT AGGAAGCTAA CCGCAGGTTG | 11040 |
| CTCAAAACAC TGTMTTGAGG TTGTGGATGA AACTGACGAA GTCAGCTCAA AACACCGTTT | 11100 |
| TGAGGTGGCA GATAGAACTG ACGAAGTCAG TAACATATAT ATACGGTAAG GCGACGCTGA | 11160 |
| CGTGGTTTGA AGAGATTTTC GAAGAGTATT AAGCTAGTTT TTAGGTTTGG CTCTTATACT | 11220 |
| AGAGTCATCA AAAAGAAACG AGGACTCTCA TATGACAGTA ACGATTAAAG TAAATTACCA | 11280 |
| AACCACTTTC CAAAAGAAGG AAGCAAAAAA CTAGTATAAA CAGAAGAGAG AGCGAAATGC | 11340 |
| TCTTTTTCG TTTCTAAAAC TACTTTCAGC CCATCATCCT AAAAGTAAAG AATCTAAATT | 11400 |
| CACTTTCTAT TTACCCTTCT TTCTTGCAAT GATTACATAG ATATGCTACA GTTGTGGTAA | 11460 |
| CGATTACAAA ATAAAAGGAG CATGCTATGA AAAATCCAGC TTTGCTAGAA GAAATTAAGA | 11520 |
| CCTATAGAGG AAGGGATGAG GTTCCGGAAG ACTTTGATGA TTTCTGGGAT GGGGAAGTGA | 11580 |

963

| | |
|--|-------|
| AAAATGTTTC CACGCTTCCA TCCTACCACT TGGAGGAAAG AGATTTCCAC ATTCCTCAAG | 11640 |
| TCAAGTGCTA TGAGTTAACA TTTGAAGGAA GCAAGGAAGG AAAGGTCTAT GCACGCATTG | 11700 |
| TTCTTCCAAA GAGTGAGGAG AAGGTCCCAT TAATCTTCCA TTTTCATGGT TATATGGGAC | 11760 |
| GTGGCTGGGA CTGGGCCGAC ATGCTGGGCT TCACCGTAGC TGGTTACGGT GTTGTTCCTA | 11820 |
| TGGATGTGCG GGGCCAGTCA GGTACTCAC AAGACGGCTT GCGTTCCTCT TTAGGAAATA | 11880 |
| CCGTGAAGGG GCATATTATC CGTGGTGCTG TGGAAAGGTCG GGACCACCTC TTTTATAAGG | 11940 |
| ATGTTTATCT GGATATTTAC CAGTTGGTCG AAATTGTTGC TAGTCTGTCT CAGGTTGATG | 12000 |
| AGAAGCGTCT TTCTAGCTAT GGTGCCTCAC AAGGAGGGGC TCTAGCTCTA GTTGACGACG | 12060 |
| CGCTCAATCC TCGAATTCAG AAAACAGTTG CCATTTATCC CTTCTTGTC AACTTCAGAC | 12120 |
| GGGTGATTGA GATTGGTAAT ACTAGCGAGG CTTACGACGA ACTTTCCGT TATTTCAAGT | 12180 |
| TTACACGACCC CTTCCATGAA ACAGAGGAGG AAATCATGGC GACCCCTGCC TATATCGATG | 12240 |
| TCAAAAATCT TGCCCATCGT ATCCAAGGTG AGGTTAAGAT GATTACGGGC TTGGACGACG | 12300 |
| ATGTTTGCTA TCCCATTACC CAGTTTGCGA TTTATAATCG TCTGACCTGC GATAAACCT | 12360 |
| ATCGCATCAT GCCTGAGTAT GCTCACGAAG CCATGAATGT ATTTGTCAAT GACCAAGTCT | 12420 |
| ACAACCTGGCT CTGTGGAAGT GAGATTCCTT TTAAATATCT AAAATAAGGA GTCGACTCTA | 12480 |
| AGCACAAAAT CTTAAAAATT ACAAACACGC ATAGTATCAG GGGATTAAAG AAACCTTTATA | 12540 |
| CTATGCGTTT TATCATGGAA ATATAGTAAA ATGAAATAAG AACAGGACAA ATCGATCAGG | 12600 |
| ACAGTCAAAT CGATTTCTAA CAATGTTTGA GAAACAAATG TGTACTATTC TAGTGTCAAT | 12660 |
| CTATTATATT TATAGAATTT TTTGTTGCTA GATTGTGCAA ATTGCTTAAA ATAATTTTTT | 12720 |
| TCAGAAAGCA AAAGCCGATA CCTATCGAGT AGGGTAGTTC TTGCTATCGT CAGGCTTGTC | 12780 |
| TGTAGGTGTT AATACTTTTC AAAAATCTCT TCAAACCACG TCAGCTTCGC CTTGC | 12835 |

(2) INFORMATION FOR SEQ ID NO: 142:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 5020 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 142:

| | |
|---|-----|
| GGGGATATGA AGAACAAAAG AATATTTAAA GACTTCCAAG CTTCAAAAAT GAGTTTAAAC | 60 |
| ATTTACACAA GCCCCTTGTT AGCCTTTGTT TTTGCTTCA TAGGAGAGTT TGTGGCTTTT | 120 |

964

| | |
|--|------|
| ACTTTGTATG GTATTGGCTT GTTAGCTCTC ATCGGACTTG CTAGAAATTT TGGAGAGGCT | 180 |
| GGTCAAAATC TTGCAAGCTA CTTGCAGACC TTGCATCAGA GCTTGACGGA TAAAACAAGT | 240 |
| GACTTTCGTT TAATTTTAGG ATTAAGTGGC TTTGGTTATT CTTAACACTG TGTTTCAGATG | 300 |
| GACAAGAAAA GTTGAGAAAA GACCTATTCG AACCTTGGGA TTTTATAGAG AGAATTTCTT | 360 |
| CAGCAATCTT CTGAAAGGAT TTAGTCTAGG CCTGGCACTT TTTCTTCTGA CCTGTGTTAGG | 420 |
| TTTAGTGGTC TTAGGTCAAT ATCGTTTGA ATCCATTAC TTGAATCCTT ATTCTCTTGC | 480 |
| CTTTGTGCTC TTTACTATCC CATTTTGGAT TTTACAGGGG ACAGCAGAAG AAGTGGTGGC | 540 |
| CCGTGCTTGG CTACTTCCTC AATTGGCCTC AAGAACCAAT CTAAACTAG CTATTCTTAT | 600 |
| ATCTAGCCTG TTCTTTACCC TGCTTCATAT GGGCAATTCT GGTCTCACCC CTCTATCTCT | 660 |
| AGTAAATCTC TTTTATTCG GAGTTGCCAT GGCTCTTTAC CTTCTCAAAA CTGATACAGT | 720 |
| TTGGGGTGTG GCAGGTATTC ATGGTGCTTG GAATTTTGCT CAGGGTAATC TCTTTGGGAT | 780 |
| TTTAGTTAGT GGTCAACCGT CAGAACGCT CTGATGACCT TTTTACCACA AGGCAATCAA | 840 |
| GATTGGCTAT CAGGTGGTTC TTTTGGCATA GAAGGTTCCA TTATGACAAG TCTGGTATTA | 900 |
| CTACTGCTGA TTGTCTATCT TGCTAATAAA TTAAGAAAG AAAATGAAAG GATGTGACTT | 960 |
| CGGTCCGCTC TTTTCTTCGT GAAAATACTA TAAGTATGCT AAAATAGGAA TAGCACATGG | 1020 |
| AGAGAGGATT CTTATGATCA ATCACATTAC AGATAATCAA TTTAACTAG TATCAAAATA | 1080 |
| TCAACCATCA GGAGATCAAC CCCAAGCTAT CGAGCAGTTG GTGGATAACA TTGAGGGGGG | 1140 |
| AGAAAAAGCT CAGATTCTGA TGGGGGCGAC TGAACAGGG AAGACCTATA CTATGAGTCA | 1200 |
| GGTCATTTCT AAAGTCAATA AACCAACTCT GGTATTGGC CACAATAAAA CTCTGGCTGG | 1260 |
| TCAGCTCTAT GGGGAGTTTA AGGAATTTT CCCTGAAAAT GCAGTTGAGT ATTCGTATC | 1320 |
| CTACTATGAT TATTACCAGC CAGAGGCCTA TGTCCCTTCT AGCGATACCT ATATTGAGAA | 1380 |
| GGATAGTTCT GTCAATGACG AGATTGACAA GCTTCGCCAC TCAGCTACCT CAGCCCTTTT | 1440 |
| GGAGCGTAAT GATGTATTG TCGTGGCCTC AGTCTCTTGT ATCTATGGTT TGGGTTGCC | 1500 |
| CAAGGAATAC GCTGATAGTG TCGTTAGTCT CCGTCCTGGT CTAGAGATTT CTCGTGATAA | 1560 |
| ACTCTTGAAT GACTTGGTCG ATATTGAGT TGAACGTAAT GATATTGATT TCCAACGCGG | 1620 |
| AAGATTTGCG GTTCGTGGGG ATGTGGTAGA GATTTTCCCA GCTTCCCGAG ATGAACATGC | 1680 |
| CTTTCGAGTA GAATTTTTTG GAGACGAAAT TGACCGTATT CGTGAAGTTG AGGCTCTGAC | 1740 |
| AGGTCAGGTG TTGGGAGAAG TGGATCATTT AGCGATTTTC CCAGCGACAC ACTTTGTGAC | 1800 |
| CAATGACGAC CACATGGAAG TTGCCATTGC AAAGATTGAG GCCGAGTTGG AAGAACAATT | 1860 |
| AGCTGTCTTT GAAAAGGAAG GTAACTGCT TGAAGCCAG CGTTTGAAAC AGCGGACAGA | 1920 |

965

| | |
|--|------|
| GTATGATATC GAAATGTTGC GTGAGATGGG CTATACCAAT GGGGTGAAA ATTATTCTCG | 1980 |
| CCACATGGAT GGACGGAGCG AAGGAGAGCC TCCTTATACG CTTCTCGACT TCTTCCAG | 2040 |
| TGATTTCTTG ATTATGATTG ACGAGAGTCA TATGACCATA GGGCAAATCA AGGGCATGTA | 2100 |
| CAATGGAGAC CGTTCGCGTA AAGAAATGCT GGTAAATTAT GGTTCGGTT TGCCGTCTGC | 2160 |
| TTTGGACAAT CGTCCTCTCC GTCGGGAGGA GTTTGAGAGT CACGTTTCATC AGATTGTTTA | 2220 |
| CGTTTCAGCG ACACCTGGTG ACTATGAAAA TGAACAGACC GAGACAGTGA TTGAGCAAAT | 2280 |
| CATTCGTCCA ACGGGACTCT TGGATCCAGA GGTGGAAGTC CGTCCGACTA TGGGACAGAT | 2340 |
| TGATGACCTC TTGGGTGAAA TCAATGCCCG CGTTGAAAAA AATGAGCGTA CCTTATCAC | 2400 |
| AACCTTGACC AAGAAAATGG CAGAGGATTT GACCGACTAC TTCAAGGAAA TGGGTATCAA | 2460 |
| GGTCAAGTAC ATGCACTCGG ATATCAAGAC CTTGGAACGG ACGGAGATTA TCCGTGACCT | 2520 |
| GCGCTTGGGT GTCTTTGATG TCTTGGTCGG AATTAACTTG CTCCGTGAAG GAATTGACGT | 2580 |
| TCCTGAAGTG AGCCTCGTAG CTATTCTCGA TGCTGACAAG GAAGTTTCC TTCGCAACGA | 2640 |
| ACGTGGACTC ATCCAGACCA TTGGACGTGC TGCACGTAAT AGCGAAGGTC ATGTTATCAT | 2700 |
| GTATGCGGAC ACGGTTACCC AGTCTATGCA ACGTGCTATC GATGAAACTG CCCGCCGTCG | 2760 |
| CAAAATCCAG ATGGCCTATA ATGAAGAACA TGGTATCGTT CCACAAACCA TCAAGAAAGA | 2820 |
| AATCCGTGAC TTGATTGCTG TGACCAAGGC AGTTGCTAAG GAAGAAGACA AGGAAGTCGA | 2880 |
| TATCAATAGC CTCAACAAAC AAGAGCGCAA AGAACTAGTC AAAAAGCTTG AGAAACAAAT | 2940 |
| GCAAGAAGCA GTTGAAGTGC TTGACTTTGA ACTAGCAGCT CAGATTTCGTG ATATGATGCT | 3000 |
| GGAAGTCAAG GCCTTGGATT AGGGGAATAG TATGATTTAT TTAAGAAAGT TAAAGAAAGA | 3060 |
| AGATTTGATG TCTTTATGGG AAATGGCTTA TTCACAGCTT AATCCAGTTT GGAAACAGTA | 3120 |
| TGATGCTCCC TATTATGATG ATTATCAGTA TTTTCAAAT TTTAAAGAAT TCGAACTACA | 3180 |
| AAAATCAGAA TCCATTTTAA GCAACTCAAA TCGCCTTGGT ATTTTGTGTG ATGATAAACT | 3240 |
| AGTTGGGACT GTTTCGCGTT ATTGGGTATG TAAAGAAACA AGATGGATGG AATTGGGAAT | 3300 |
| TGGTATTTAT GATAAAAAAT TCTGGAACAC TGGTATTGGG AAAGTTGCTA TGTTCAGTGC | 3360 |
| GATAGATAGG ACGTTTCAGG ATTACTTGA GTTGGAGCAT CTGGGTTTGA CAACTTGGTC | 3420 |
| AGGAAATATT GGTATGATGA AACTTGCTGA AAAATTAAGA ATGAAAAAAG AAGCTCATAT | 3480 |
| TCCAAAAGTT CGTTATTATC AAGGTAAATA TTTTGACAGT ATTAAATATG GTATTTTGAG | 3540 |
| AGAAGACTGG GAGAAAATAA ATGACGGTTA TTATCAAATC AATGGAAACT CCTGAAGAGA | 3600 |
| TAGAAGGTAA ATCCTTCGTT CACTGGCAAA CGTGGAGAGA GGCTTATGAT GACCTTTTGC | 3660 |

966

| | |
|---|------|
| CTGCGGAATT TCAGGAGACA ATGACATTAG AAAGATGTCG ACTCTTTAGT CAAAAGTATC | 3720 |
| CAGAAAATAC ATTGATTGCG ATGGATGGTG TGAAGATAGT TGGTTTTATA AGTTATGGCA | 3780 |
| ACTGTCGTGA TGAGACTATT CAAGCTGGTG AAATTATTGC TTTATATGTT TTAAGAACT | 3840 |
| ATTATGGAAA AGGAATCGCA CAAAAGTTAG TGAAAGCAGC TTTGACTGAT CTTAATCATT | 3900 |
| TTTCTGAAAT TTTCTTATGG GTATTGAAAG ATAACAAGCG CGCCATTGCT TTCTATCAAA | 3960 |
| AAATGGGTTT TACTTTTGAT GGACAAGAAA AAATACTTGA ACTTGGAAAG CCTATAAAGG | 4020 |
| AAAAACGGAT GGTATTCTAT TCTAAATAAT TCTCAAAAGT AAAAGCTAAT ATGGTACCAA | 4080 |
| GTCTGAAAAT TTAATAAATT AGAAAGCGAG TAAATTTATG TCCCGTTCCC AATTAACAAT | 4140 |
| TTTAACAAAT ATCTGTCTGA TTGAAGACCT CGAACTCAG CGCGTGGTGA TGCAGTATCG | 4200 |
| CGCCCCTGAA ACAATCGCT GGTCTGGTTA TGCCTTTCCT GGAGGTCATG TAGAAAATGA | 4260 |
| TGAGGCTTTT GCGGAGTCTG TCATTCTGTA AATCTACGAA GAAACAGGGT TGACTATCCA | 4320 |
| AAATCCTCAA CTTGTCGGCA TTAAAAATTG GCCACTAGAT ACAGGTGGGC GCTATATTGT | 4380 |
| CATTTGTTAT AAGGCGACTG AGTTCCTCTG TACCCTTCAA TCTTCAGAAG AGGGAGAAGT | 4440 |
| TTCTTGGGTG CAAAAGACC AGATTCCAAA CTAAATCTG GCCTATGATA TGCTACCATT | 4500 |
| GATGGAAATG ATGGAAGCTC CCGACAAGTC AGAGTTTTC TACCCTCGCC GTACAGAAGA | 4560 |
| CGATTGGGAA AAGAAAATCT TCTAGTCTTT TACTAAATAA CCTAGCTGAT CCAAGGCCTC | 4620 |
| CTCGATATAG TGGAGGTCTT GTTGTGTCTC GGCTTCAACT AGGTGATAAT GAATACCATC | 4680 |
| TGTTAACTCA GAAATTGGCT TAAAGTCAGA ACGTTCAACT TGTTCTAGAA AATGTTGCAC | 4740 |
| GTCGCGGCGA CAGGTCAGTT TTAGTAAGGT TTCAATCTCT CCATAAACAG GATGATCAAT | 4800 |
| CAAGATATTT TGAACGCGAC CACCATATC TACGATAGCA AGTAATTCTC GTCCAATTTC | 4860 |
| TTCAACTTCA TGCTTGACCT TAAATAATTT GTGATGATAA GTATTGTCAT TAGCATCTTT | 4920 |
| ATAGATATAA CCACGATTGG TAGATAGAAT TGGAGATCCA TCAGCTCTTA AAATTGCAAT | 4980 |
| ATCTTGAACA ATAACTTGTC GAGTGACATG AAAGTGCTCA | 5020 |

(2) INFORMATION FOR SEQ ID NO: 143:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 4965 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 143:

| | |
|---|----|
| AAAAAGTGGC AATCCATTGA TTGGCCACTT CATTTAGAGA ATTATCGTCT CGCCCTTGAA | 60 |
|---|----|

967

| | |
|---|------|
| GAAGAAGGTC GTGTAGTACT TGAGTTACTG CTATCGCTAG AACTACTACT TGAAGTCTG | 120 |
| GAGCTGGATG GAGTTGGTAG ACTCCCCACA ATACTAGACC AAGCATTCTG ATAATCCGCA | 180 |
| TCACTTCCGC CAATAGCAAA GCGATAACTT GTCGCTGGCG CTCCTGACTT ATTAGCCCAA | 240 |
| TAGCTGGTAA CAGTCGAACC TGTGACCTCT ACTTCTTTTC CTTCAACAGA AACCTTCTCT | 300 |
| GGTTTTTGAC CTGTTGATTT CAAGACTTCC GATTTCACTA CACTAGGATC TAAAGCAAAG | 360 |
| CGCTCGTTCC CCCAAATGCT TGGGGAAGCT TGCTGAATCG CATTTACCAG ATGAGCCATG | 420 |
| TAATTAGAGT TATTAGAATA ACCTGCTCTA CGTGACAATG AATGATTATC ATCATGCCCA | 480 |
| ATCCAGCCAC CTAGGGTTAA TCTAGGTGTC GAAAGCATGA GCCACATATT TTCGTCTTGG | 540 |
| TTGGTTGTAC CAGTCTTCCC AATCCAATCT GCATTAGCCA GAGTAGGATT TAAAGAAGTC | 600 |
| AGGTTAGACT TGAAGGTTGT TGTACACGA GAGGATAGAA CTTCTCGTAG CAATCCCTGC | 660 |
| ATAATCGTCG CAGTAGCTTT TGAATAGACT TGAACCGGTT TATCCTGATA CTCATACACC | 720 |
| ACTCTACCAT CTGCTGCTTC AATCTTTGAA ATCACATGCT TCTGATGATA AACTCCATTA | 780 |
| TTAGCTAAGG TCTGATAGCC ATTGGTATGC TGGGCAACTG TGACTTCAAT ACCACCACCC | 840 |
| ATTGGCAAGC TCTCAATACC GTACTCAGGA ATCTCGTAAC CCATCTTTTC CATATAACCC | 900 |
| TTGACATCAA CACCCTTTTC ACGGAGCATA CGATAGGTCC AGTAAGCAGG GATATTCCAT | 960 |
| GAATAGTTCA GAGCTTCTCC CAAGGTCATC ATTCCTGTTT CCTTGCTATT AGCATACATA | 1020 |
| ATCGGATTGC CATTAGCAAA GTTTGTTGGA TAGTTAGATA GAATCGTTTC ACTTCCCATC | 1080 |
| AAGCCCTGGT CAATAGCAAT ACCGTAGGCC AGCAAGGGCT TGGTAGTAGA AGCTGGCGAA | 1140 |
| CGTTTGATAT CAAAGGCATG ATTATTTTGA TTTTCTTGAT AATTACGACC ACCTACAAAG | 1200 |
| CCTAGAATAG CACCTGTTTG GTTATCCATC AAGACATTCC CTACTTCTAC ACGACCTGTT | 1260 |
| CCATCGTCTA AAAGATAGCC ATAATCAGCA ACCGCACTTT GCATGGCAGA ATGAATTTTC | 1320 |
| TGATCTATGG TAGTAGTAAT CTTATAACCA CCATTTTCAA TTTCTTGGC TGCCAAATCT | 1380 |
| CGATAAACT TCTGAGTTGC CTCATTTTTC AACTCCTTAG CGGAGACATT GTCTCTCTGA | 1440 |
| GCTAGATAGT CATAACATG TTCTTGAGCT TCTGCCAAAG TTGTAAAGTA TAAATAGTCT | 1500 |
| CGTGAAATTC CTGTAACCGT GCCCGATGGT AAAAAGTCCT GTTTAAGGTC ATAATCCTTG | 1560 |
| TACTGAGAAT ACTCGTCTTT GCTTAATGCA CCTGTACGAT ACATACTGTA AAGAACTGCC | 1620 |
| TTAGCCCGTC TTAAGCCAAT TTCTAGGTCT TCATCACTCT TCAACTCCCC AGTATTTTCA | 1680 |
| TAAGGAGAGT AAGTAATGGG ACTCTGTGGA AGTCCTGCTA AAAATGCTGC TTGAGGAACA | 1740 |
| GTCAACTGAC TGGCATCTAC ACCGAAAATT CCCTCAGCTG CTTGCCGAGC CCCTGCAATA | 1800 |

968

| | |
|---|------|
| TTCTGTCCCT TATTATTTTCG GCCAAAGGGA GCCACATTGA GATAGGTCGT TAAAATCTCA | 1860 |
| TCTTTATTCA TGGCGCGTTC CAAGGCAAGA GCATCCACAA TCTCTGCCGC CTTACGAGCC | 1920 |
| AAGGTCGGCG CATCCCCAAC CACCTGCTGT TTAATTAGTT GCTGGGTCAA GGTGAACCC | 1980 |
| CCACTAGAGG AACCCAAACC TACAAATTTC CCCAAGGTCG CACGAATCAC CGCCTTGGGT | 2040 |
| ACTACACCCT TATGTTCTTT AAAGTGTTC TCTTCTGTCG CAATGATAGC CTTCTTCAGA | 2100 |
| TTTTCCGAAA TTGCTCAGA TGAGATAGAA GTGCGCAACA AATCACTCTC TATGGAAGCA | 2160 |
| ATCACCGTCC CGTCCGAATA GGTAATCTCT GAAATAGAAG AGATGTCCTT GACCTGATTC | 2220 |
| ACCAATTCTT CTGCTGAGG CACCCGAACC TTGTCAAATA AGGCCACTCC GTATCCCAAA | 2280 |
| GCAATCCCAG CTCCCAACAT TCCTCCTAGA AAACCGAGTA CAAAGAGTAA GTTAAATAAG | 2340 |
| GCTTTTATAC TCAGTAAAAT AGCTGGGAAA ATGACTGACT TATCTAAGGT TTTAGATTTT | 2400 |
| TTGGTACTTG AACCTTTCTT GCCAGGTCTA GCTGATTTTT TATTTTTTTG TTTTTGCTGG | 2460 |
| AAAAATTCCA GCATTTTTCG TTTTAATTCA TTAAATTGAT TTTGCATGGA TTTCTCACT | 2520 |
| TTATCTATTA TACCACAAA GGGAAATTT CAATAAAATA GCCACTTTCT TCCCTATTCT | 2580 |
| GCTAGGCTAT TGCCCAAGTT TGTGATACAA TAGGTAGAAA CAATAATTTT AAAAAGGAGA | 2640 |
| AAAAACACAT GCACATTTTT GATGAGCTAA AAGAGCGTGG TTTGATATTT CAAACGACTG | 2700 |
| ATGAAGAAGC TTTGCGTAAA GCCCTAGAAG AAGGTCAAGT TTCTTATTAT ACTGGCTACG | 2760 |
| ATCCAACGTC TGACAGCCTT CACCTAGGCC ACCTTGTCGC AATCTTGACA AGTCGTCGCT | 2820 |
| TGCAACTAGC AGGTCACAAA CCTTATGCGC TCGTTGGCGG TGCTACAGGT CTCATCGGAG | 2880 |
| ATCCGTCCTT CAAAGATGCT GAACGTAGTC TCCAAACAAA AGACACAGTA GATGGCTGGG | 2940 |
| TCAAGTCTAT CCAAGGACAA CTTTCTCGTT TTCTTGACTT TGAAAATGGC GAAAACAAGG | 3000 |
| CTGTCATGGT CAACAACAC GACTGGTTTG GCAGCATCAG CTTCATTGAC TTCCTCCGTG | 3060 |
| ATATTGGAAA ATACTTCACG GTCAACTACA TGATGAGTAA GGAATCTGTT AAAAACGGA | 3120 |
| TCGAAACAGG AATTTCCTTAC ACTGAGTTTCG CTTACCAAAT CATGCAAGGG TATGACTTCT | 3180 |
| TCGTCCTTAA CCAAGACCAT AATGTCACCTC TTCAAATCGG TGGTTCCTGAC CAGTGGGGAA | 3240 |
| ATATGACAGC TGGTACCGAA TTGCTTCGTC GTAAGGCGGA CAAGACTGGT CACGTTATCA | 3300 |
| CTGTTCCACT AATCACAGAT GCAACTGGTA AGAAATTTGG TAAATCAGAA GGAAATGCCG | 3360 |
| TCTGGCTCAA TCCCGAAAAG ACTTCTCCAT ACGAAATGTA CCAATTCTGG ATGAACGTGA | 3420 |
| TGGACGCTGA CGCTGTTTCG TTCTTGAAAA TCTTTACTTT CTTGTCACCT GATGAGATTG | 3480 |
| AAGATATTTCG TAAACAATTT GAAGCAGCGC CACACGAACG CTTGGCTCAA AAAGTCTTGG | 3540 |
| CTCGTGAAGT TGTTACACTT GTTCACGGAG AAGAAGCCTA CAAAGAAGCA CTTAACATCA | 3600 |

969

| | |
|---|------|
| CTGAGCAACT CTTTGCAGGA AACATCAAAA ACCTTTCTGT CAAAGAGCTC AAACAAGGAC | 3660 |
| TTCGTGGTGT GCCCAACTAC CAAGTACAGG CAGACGAAAA CAACAATATC GTGGAAGTGC | 3720 |
| TCGTCTCATC TGGTATAGTT AACTCAAAAC GCCAAGCCCG TGAAGACGTC CAAAACGGAG | 3780 |
| CCATCTACGT AAACGGCGAC CGCATCCAAG AGCTTGAATA TGTCTTGAGT GACGCTGATA | 3840 |
| AGTTAGAGAA TGAAGTGAAT GTTATCCGTC GTGGGAAGAA AAAATACTTT GTATTGACTT | 3900 |
| ACTAAACTAT TCAACATTTA TCTATAACA AAGGAGTTAA CCTCGAGAAA GGTAAGTCTT | 3960 |
| TTTGTCTGTT AATAACTCTC ATCTATCTAT TTTTAATAGA CAGGCTACGC AGGACAATGC | 4020 |
| GCAAGGTGTT TAGATTATGT AAGATAGAGA GATTGAAGG ACTGAACCAA TTAATAAAGC | 4080 |
| CAAAGCCAAT CAAACTACTA TTTACGACAA CGGTATCCTG AATATTTTTC TTGATGAGTG | 4140 |
| TTTGCAAAGA TGATGATAAC GAATCCAAT CTTGGAAGAA ATCCAAACGA TTATCTAACA | 4200 |
| ATAAGATATC ACTCATCTGC TTAGAAATAT CTGCACTCTC ATTCATCACC ACACCGATAT | 4260 |
| CTGATAGAGT TAAAGCCGCT GAGTCATTCA ATCCATCTCC AACCATCAAA ATAGTGTGAC | 4320 |
| CTGCTTTCTG CAGTTTCTCT ACTAACTCAA ATTTCCCATC AGGTTTCAAG TCTGTATAGA | 4380 |
| CCTGATCAAA GGGCAAATCT TTGACTAATT CCTCTGTCTT AATCAAGGTG TCTCCTGTTG | 4440 |
| CCAGAATCAA TTTTTCCTCC TGTGCCTTAA GTTATCCAA GGCTGTTTTT GCTTCTTTTC | 4500 |
| TCAAAGGAGT ATGAATGCAG AACATTCCAA TCAATTCAAT TTGATAAGCC AAGAATAAGA | 4560 |
| GATTGTAGTG ACTCTGTAC TCTTCAATTA AAGCATTTTG TTCTGAACTG ATATGAATCT | 4620 |
| GCTCATCCTG CATCAAGACA TAATCCCAA TAAGAACTGG TTGGCCATCT ATATGAGATT | 4680 |
| TGATCCCTTT GCTTGCGATA TATTGGAGTT TCCCATGCAT TTCCTCATGT TCAATTCCCT | 4740 |
| CTATCTCAGC TTGCTTGACG ATGGCATTAG CAATAGGATG ATAAATGTGT TCCTCAAGAC | 4800 |
| AGGCACTGAT TCTGAGAATA TCTTCTCAC TATAGTCTCC AAAAGGTAAC ACCTTTTCAA | 4860 |
| CTATAGGATA ACTAGTTGTG ATTGTTCTTG TCTTATCAAA CAAGAAAGTA TCAACTTCCA | 4920 |
| GATATTTCTC CCTGTTGTGG CCTCTGGCTG TCATCTCTGT GCTGG | 4965 |

(2) INFORMATION FOR SEQ ID NO: 144:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 3232 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 144:

970

| | |
|--|------|
| CAGGGGCGTA TTACGTGACA ATTCAATGTA GGCTGTCGCT ACTTGCGCCA AAACAAGGAT | 60 |
| TCGATAATGT CGGATGATAC TAACGATTAA ACCGAGCAGA AAGGATCCCA AAATCCCCCA | 120 |
| AACTGCAATA TGCAAGGTCA GAAAGAATGC CTTTGTATAT AGTGGTAGAT ATTGTTCAAC | 180 |
| AATGGATCAA TCCAAAAATA GAACCTCCCA TCTAGAAATA ATACAGTTAT TGTAGCACTT | 240 |
| AAAATCTTCT TTGGATAATA TCTATTTTTT ATTGCCGTTA TAAGGATTTT TATCATAGAC | 300 |
| ATAAAATTTC TGAAATTTCC AAACAAAATA TTTTAAAAGT TTTGAAAAAG AGTTAAGATA | 360 |
| TTTTTGTAAAT ACACAAAGTA AACGCTTACT TATTAAGGAG GACATTTTAT GTCATACAAA | 420 |
| ACAAGCAATG CAGAAGGTCA TGTAAGTTTC ATCAATACCT ATGATTTGGA GCCAATGGCG | 480 |
| CAACAAGTTA TTCCTAAAGC AGCATTGGC TATATCGCTA GTGGGCGGG AGATACTTTC | 540 |
| ACTTCTTTCC AGTGATTTTA GCGTCAGGT CTTTTAGTT TTTAAAGATT ATCCGTGAAT | 600 |
| TTCTTGCTTA TTTATGATAA AATGGGAGTG TCGCAAAAAA TGACTCATCG TATTCAATTT | 660 |
| TGAGTAAAC TAGGAGGATC CCATGTCTAC AGAACATATG GAAGAACTAA ATGACCAGCA | 720 |
| GATCGTTCCG CGTGAAAAA TGCTGCGCT CCGCAACAA GGAATCGATC CTTTCGGAAA | 780 |
| ACGTTTGA CGTACTGCAA ATTCACAAGA ATTAAGAT AAATATGCCA ACCTCGATAA | 840 |
| AGAACAATTA CACGATAAAA ACGAAACAGC TACTATCGCA GGACGCTTGA TAACCAACG | 900 |
| TGGTAAAGGA AAAGTTGGTT TTGCCCACCT TCAAGACCGC GAAGGCCAGA TTCAGATCTA | 960 |
| CGTTCGTAAG GATGCTGTCG GTGAAGAAA CTACGAAATC TTCAAAAAG CAGACCTTGG | 1020 |
| TGACTTCTT GGTGTCGAAG GTGAAGTGAT GCGTACGGAT ATGGGAGAAC TCTCTATCAA | 1080 |
| GGCAACCCAC ATCACACACT TGTCTAAGGC TCTCGTCCT CTTCCTGAGA AATTCCATGG | 1140 |
| TTTGACAGAC GTTGAAACAA TTTACCGTAA ACGTTACCTT GACTTGATTT CTAATCGTGA | 1200 |
| AAGCTTTGAA CGCTTTGTCA CTCGTTCAAA AATCATCTCT GAAATCCGTC GTTACCTTGA | 1260 |
| CCAAAAAGGA TTCCTTGAAG TGGAACACC TGTTCTTCAT AATGAAGCCG GTGGTGCTGC | 1320 |
| TGCCCCGCCA TTTATCACCC ACCACAATGC CAAAACATT GACATGGTGC TTCGTATCGC | 1380 |
| GACTGAGCTT CACTPAAAC GCCTTATCGT GGGTGGTATG GAACGTGTCT ATGAAATTGG | 1440 |
| CCGTATCTTC CGTAACGAAG GAATGGACGC TACTCATAAC CCTGAGTTCA CTTCTATCGA | 1500 |
| AGTTTACCAA GCTTATGCAG ACTTCCAAGA CATCATGGAC TTGACTGAAG GCATTATCCA | 1560 |
| ACACGCTGCT AAATCAGTCA AAGGTGATGG CCCAGTCAAC TACCAAGGTA CTGAAATCAA | 1620 |
| GATTAAACGAA CCATTTAAGC GTGTTTATAT GGTGGATGCT ATCAGAGAAA TTACTGGTGT | 1680 |
| CGATTTCTGG CAAGACATGA CTTTGAAGA AGCTAAAGCT ATCGCTGCTG AGAAGAAAGT | 1740 |
| TCCAGTTGAG AAACACTACA CTGAGGTTGG TCACATCATC AATGCCTTCT TTGAAGAGTT | 1800 |

971

| | |
|---|------|
| TGTTGAAGAA ACTTTAATCC AACCAACCTT TGTCTATGGA CATCCAGTAG CTGTATCTCC | 1860 |
| ACTCGCTAAG AAAAATCCTG AAGACCAACG CTTTACTGAC CGTTTCGAGC TCTTTATCAT | 1920 |
| GACTAAGGAG TACGGTAATG CCTTTACTGA GTTGAACGAC CCAATCGACC AACTTAGCCG | 1980 |
| TTTTGAAGCC CAAGCTAAAG CCAAAGAACT TGGTGATGAT GAAGCGACAG GAATCGACTA | 2040 |
| TGACTACATT GAAGCTCTTG AATACGGTAT GCCACCAACA GGTGGTTTGG GAATCGGTAT | 2100 |
| CGACCGTCTC TGCATGCTCC TCACTGATAC AACAACTATC CGTGATGTAT TGCTCTTCCC | 2160 |
| AACAATGAAA TAAATTCCTA TCCTCTGGGT CTTATCAGAG GATTTTTTGA TTCAAAAAGA | 2220 |
| GACTGAATTT AAGGAGAAAA TGAAGTGTAG TATATTGAAA TTGAAATAGT AACTTTTGAT | 2280 |
| TTCTAAGACA TTGTTAGAAA TTGGTTTAAA TTCCCTAAGC AATTTGTGCA TGTTTTATTT | 2340 |
| CATTTTACGA TAGTACGCTG AAACCTTTCA AAAAGTACTA GAAATTGACT TGGATTCCCC | 2400 |
| AATTGATTTG TTCAGATTCA CTATAAATAA AAAATTAATA AGTGGGATAG GAAGTTAGCG | 2460 |
| TCAACTAGGA TAGTATCTTG CTTAAACAGT ATATATGGGA TTGATATAAG TCCATAGGTC | 2520 |
| CTATTAGAGG ATGTTCTGGT GTCTTATTCA CTTGTTTTTT ATAGTATTAG TAGATAGAAT | 2580 |
| CAGCAATAA AAACCCAAAT CATTCATACC TCTCTCACT AGATGTAACT TACAAAACCC | 2640 |
| CTGACCTCAT GAGCCACTTT CTTCTCCTC ATGAGGTCAG TTTTACTTTC TGCTGTTCCA | 2700 |
| GTATCGTTTT TCCTCGCTAG ATTTCTCTCA AAGGGCAGAC TCCTCCCTTG GTGCGTCACA | 2760 |
| CGATTTTTTC ATCTCGACTG TTCTTTAATG CATCATTAAC GACGCTTTTC TTCTAGGTGG | 2820 |
| TTCATAAGGA ACAGGAAGAT TCAGGTTGAC TTTTCTAATC CTAGAATAAA GTGCTGAAAA | 2880 |
| CAATTCGGAA TAGGCATAGA GACTAGACAA TTTGAGGAGC TGCTTGCGTC CTGTTCGAAC | 2940 |
| ACATTTTCCC ACCACGTGAA GAAAAAGATG GCGGAAGCGT TTGATTGTTA AAGTTTGGAA | 3000 |
| GTCACCTCCA GCTAGATGTT TGAGAAAAAG ATAGAGATTG TAGGCGATAC AGCTCATCAT | 3060 |
| CATACGAAT TCGTTTTTGA TTAAGGTTGA ACTATCCGTT TTATCGCCAA AAAATCCCTC | 3120 |
| CTTCATCTCC TTGATGAAAT TCTCGGCTTG ACCACGTCCA CGATAAAGCT GAAACTGGTC | 3180 |
| TTGGCTTGTT CCACTCGTCA TATTTGTAAC GAGAGAAATA ACATCGTAGA AC | 3232 |

(2) INFORMATION FOR SEQ ID NO: 145:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10711 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

972

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 145:

| | |
|--|------|
| CCGGAGAAAA TGATGAAAAG TTCAAAAC TA TTTGCCCTTG CGGGCGTGAC ATTATTGGCG | 60 |
| GCGACTACTT TAGCTGCATG CTCTGGATCA GGTTCAGCA CTAAGGTGA GAAGACATTC | 120 |
| TCATACATTT ATGAGACAGA CCCTGATAAC CTCAACTATT TGACAACTGC TAAGGCTGCG | 180 |
| ACACAAATAT TACCAGTAAC GTGGTTGATG GTTTGCTAGA AAATGATCGC TACGGGAACT | 240 |
| TTGTGCCGTC TATGGCTGAG GATTGGTCTG TATCCAAGGA TGGATTGACT TACACTTATA | 300 |
| CTATCCGTAA GGATGCAAAA TGGTATACTT CTGAAGGTGA AGAATACGCG GCAGTCAAAG | 360 |
| CTCAAGACTT TGTAACAGGA TTAAAATATG CTGCTGATAA AAAATCAGAT GCTCTTTACC | 420 |
| TTGTTCAGA ATCAATCAAA GGGTTGGATG CCTATGTAAA AGGGGAAATC AAAGATTTC | 480 |
| CACAAGTAGG AATTAAGGCT CTGGATGAAC AGACAGTTCA GTACACTTTG AACAAACCAG | 540 |
| AAAGCTTCTG GAATTCTAAG ACAACCATGG GTGTGCTTGC GCCAGTTAAT GAAGAGTTTT | 600 |
| TGAATTCAAA AGGAGATGAT TTTGCCAAAG CTACGGATCC AAGTAGTCTC TTGTATAACG | 660 |
| GTCCTTATTT GTTGAAATCC ATTTGTGACCA AATCCTCTGT TGAATTTGCG AAAAATCCGA | 720 |
| ACTACTGGGA TAAGGACAAT GTGCATGTTG ACAAAGTTAA ATGTGCATTC TGGGATGGTC | 780 |
| AAGATACCAG CAAACCTGCA GAAAACCTTA AAGATGGTAG CCTTACAGCA GCTCGTCTCT | 840 |
| ATCCAACAAG TGCAAGTTTC GCAGAACTTG AGAAGAGTAT GAAGGACAAT ATTGTCTATA | 900 |
| CTCAACAAGA CTCTATTACG TATCTAGTTG GTACAAATAT TGACCGTCAG TCCTATAAAT | 960 |
| ACACATCTAA GACCAGCGAC GAACAAAAGG CATCGACTAA AAAGGCTCTC TTAAACAAGG | 1020 |
| ATTTCCGTC GGTATTGCC TTTGGATTG ACCGTACAGC CTATGCCTCT CAGTTGAATG | 1080 |
| GACAACTGG AGCAAGTAAA ATCTTGCGTA ATCTCTTTGT GCCACCAACA TTTGTTCAAG | 1140 |
| CAGATGGTAA AAACCTTGGC GATATGGTCA AAGAGAAATT GGTCATTAT GGGGATGAAT | 1200 |
| GGAAGGATGT TAATCTTGCA GATTCTCAGG ATGGTCTTTA CAATCCAGAA AAAGCCAAGG | 1260 |
| CTGAATTTGC TAAAGCTAAA TCAGCCTTAC AAGCAGAAG AGTCCAATTC CCAATTCATT | 1320 |
| TGGATATGCC AGTTGACCAA ACAGCAACTA CAAAAGTTCA GCGCGTCCAA TCTATGAAAC | 1380 |
| AATCCTTGGA AGCAACTTTA GGAGCTGATA ATGTCATTAT TGATATTCAA CAACTACAAA | 1440 |
| AAGACGAAGT AAACAATATT ACATATTTTG CTGAAAATGC TGCTGGCGAA GACTGGGATT | 1500 |
| TATCAGATAA TGTCGGTTGG GGTCCAGACT TTGCCGATCC ATCAACCTAC CTTGATATTA | 1560 |
| TCAAACCTTC TGTAAGAGAA AGTACTAAAA CATATTTAGG GTTTGACTCA GGGGAAGATA | 1620 |
| ATGTAGCTGC TAAAAAGTA GGTCTATATG ACTACGAAAA ATTGGTTACT GAGGCTGGTG | 1680 |
| ATGAGACTAC AGATGTTGCT AAACGCTATG ATAAATACGC TGCAGCCCAA GCTTGGTTGA | 1740 |

973

| | |
|---|------|
| CAGATAGTGC TTTGATTATT CCAACTACAT CTCGTACAGG GCGTCCAATC TTGTCTAAGA | 1800 |
| TGGTACCATT TACAATACCA TTTGCATTGT CAGGAAATAA AGGTACAAGT GAACCAGTCT | 1860 |
| TGTATAAATA CTTGGAACCT CAAGACAAGG CAGTCACTGT AGATGAATAC CAAAAGCTC | 1920 |
| AGGAAAAATG GATGAAAGAA AAAGAAGAGT CTAATAAAAA GGCTCAAGAA GATCTCGCAA | 1980 |
| AACATGTGAA ATAACTGTTG CAAAATATAA GAAAGGATTT AGTATTTCCC TTGAATGCTG | 2040 |
| AATCCTTTTT TACATTTGTA AAGAAAGATT CTAAATGTA CGGACCCCA AAAGTTGGAG | 2100 |
| CCTCTTTTTG TCAGAATAGA GAAAATTTTT GTTAATTTTA CTTGTTTCCT ATTGCTTTCT | 2160 |
| CAGCTATTAT TTGTTATATT AAAAGTATAA TTATTTTTA TTTATCAGAG TTAAGCATTG | 2220 |
| CACCTTCAGA GGAAGGAGTA TTTTTTAAAA AGAAAATGTA AACGTTTGCT CAAAATGAA | 2280 |
| AGGATTTAGA AGTTTATGAA TAAAGGATTA TTTGAAAAAC GTTGTAATA TAGTATTCGG | 2340 |
| AAATTTTCAT TAGGTGTTGC TTCTGTTATG ATTGGAGCTG CATTCTTTGG GACAAGTCCG | 2400 |
| GTTCTTGCAG ATAGCGTGCA GTCTGGTTCC ACGCGAACT TACCAGCTGA TTTAGCTACT | 2460 |
| GCTCTTGCAA CAGCAAAAGA GAATGATGGG CGTGATTTTG AAGCGCTAA GGTGGGAGAA | 2520 |
| GACCAAGGTT CTCCAGAAGT TACAGATGGA CCTAAGACAG AAGAAGAACT ATTAGCACTT | 2580 |
| GAAAAAGAAA AACCGGCTGA AGAAAAACCA AAAGAGGATA AACCTGCAGC TGCTAAACCT | 2640 |
| GAAACACCTA AGACGGTAAC CCCTGAATGG CAAACGGTAG CGAATAAAGA GCAACAGGGA | 2700 |
| ACAGTCACTA TCCGAGAAGA AAAAGGTGTC CGCTACAACC AACTATCCTC AACTGCTCAA | 2760 |
| AATGATAACG CAGGCAAAAC AGCCCTGTTT GAAAAGAAGG GCTTGACCGT TGATGCCAAT | 2820 |
| GGAAATGCAA CTGTTGATTT AACCTTCAAA GATGATTCTG AAAAGGGCAA ATCACGCTTT | 2880 |
| GGTGTCTTTT TGAAATTTAA AGATACCAAG AATAATGTTT TTGTCGGTTA TGACAAGGAT | 2940 |
| GGCTGGTTCT GGGAGTATAA ATCTCCAACA ACTAGCACTT GGTATAGAGG TAGTCGTGTT | 3000 |
| GCTGCTCCTG AAACAGGATC AACAAACCGT CTCTCTATCA CTCTCAAGTC AGACGGTCAG | 3060 |
| CTAAATGCCA GCAATAATGA TGTCAATCTC TTTGACACAG TGACTCTACC AGCTGCGGTC | 3120 |
| AATGACCATC TTAATAATGA GAAGAAGATT CTCTCAAGG CGGGCTCTTA TGACGATGAG | 3180 |
| CGAACAGTTG TTAGCGTTAA AACGGATAAC CAAGAGGGGG TAAAAACAGA GGATACCCCT | 3240 |
| GCTGAAAAAG AAACAGGTCC TGAAGTTGAT GATAGCAAGG TGACTTATGA CACGATTCAG | 3300 |
| TCTAAGGTCC TCAAAGCAGT GATTGACCAA GCCTTCCTC GTGTCAAGGA ATACAGCTTG | 3360 |
| AACGGGCATA CTTTGCCAGG ACAGGTGCAA CAGTTCAACC AAGTCTTTAT CAATAACCAC | 3420 |
| CGAATCACCC CTGAAGTCAC TTATAAGAAA ATCAATGAGA CAACAGCAGA GACTTTGATG | 3480 |

974

| | |
|---|------|
| AAGCTTCGCG ATGATGCTCA CTTAATCAAT GCGGAAATGA CAGTACGCTT GCAAGTTGTA | 3540 |
| GACAATCAAT TGCACPTTGA TGTGACTAAG ATTGTCAACC ACAATCAAGT CACTCCAGGT | 3600 |
| CAAAAGATTG ATGACGAAAG CAAACTACTT TCTTCTATTA GTTTCCTCGG CAATGCTTTA | 3660 |
| GTCTCTGPTT CTAGTAATCA AACTGGTGCT AAGTTTGATG GGGCAACCAT GTCAAACAAT | 3720 |
| ACGCATGTCA GCGGAGATGA TCATATCGAT GTAACCAATC CAATGAAGGA TTTGGCTAAG | 3780 |
| GGTTACATGT ATGGATTGTG TTCTACAGAT AAGCTTGCTG CTGGTGTTTG GAGTAACTCT | 3840 |
| CAAAACAGCT ATGGTGGTGG TTCGAATGAC TGGACTCGTT TGACAGCTTA TAAAGAAACA | 3900 |
| GTCGGAAATG CCAACTATGT AGGAATCCAC AGCTCTGAAT GGCAATGGGA AAAAGCTTAT | 3960 |
| AAGGGCATTG TTTTCCAGA ATACACGAAG GAACTTCCAA GTGCTAAGGT TGTATCACT | 4020 |
| GAAGATGCCA ATGCAGACAA GAACGTTGAT TGGCAAGATG GTGCCATTGC TTATCGTAGC | 4080 |
| ATTATGAACA ATCCTCAAGG TTGGGAAAAA GTTAAGGATA TCACAGCTTA CCGTATCGCG | 4140 |
| ATGAACTTTG GTTCTCAAGC AAAAAACCA TTCCTTATGA CCTTGGATGG TATCAAGAAA | 4200 |
| ATCAATCTCC ATACAGATGG TCTTGGGCAA GGTGTTCTCC TTAAGGATA TGGTAGCGAA | 4260 |
| GGCCATGACT CTGGTCACTT GAACTATGCT GATATTGGTA AGCGTATCGG TGGTGTGCAA | 4320 |
| GACTTCAAGA CCCTAATTGA GAAGGCTAAG AAATATGGAG CTCATCTAGG TATCCACGTT | 4380 |
| AACGCTTCAG AAACCTATCC TGAGTCTAAA TACTTCAATG AAAAAATTCT CCGTAAGAAT | 4440 |
| CCAGATGGAA GCTATAGCTA TGGTTGGAAC TGGCTAGATC AAGGTATCAA CATTGATGCT | 4500 |
| GCCTATGACC TAGCTCATGG TCGTTTGGCA CGTTGGGAAG ATTTGAAGAA AAAACTTGGT | 4560 |
| GACGGTCTCG ACTTTATCTA TGTGGACGTT TGGGTAATG GTCAATCAGG TGATAACGGT | 4620 |
| GCCTGGGCTA CCCACGTTCT TGCTAAAGAA ATTAACAAAC AAGGCTGGCG CTTTGCATC | 4680 |
| GAGTGGGGCC ATGGTGGTGA GTACGACTCT ACCTTCCATC ACTGGGCAGC TGACTTGACC | 4740 |
| TACGGTGGCT ACACCAATAA AGGTATCAAC AGTGCCATCA CCCGCTTTAT CCGTAACCAC | 4800 |
| CAAAAAGATG CTTGGGTAGG GGAATACAGA AGTTATGGTG GTGCAGCCAA CTATCCACTG | 4860 |
| CTAGGTGGCT ACAGCATGAA AGACTTTGAA GGCTGGCAGG GAAGAAGTGA CTACAATGGC | 4920 |
| TATGTAACCA ACTTATTTGC CCATGACGTC ATGACTAAGT ACTTCCAACA CTTCACTGTA | 4980 |
| AGTAAATGGG AAAATGGTAC ACCGGTGACT ATGACCGATA ACGGTAGCAC CTATAAATGG | 5040 |
| ACTCCAGAAA TGCAGTGGA ATTGGTAGAT GCTGACAATA ATAAAGTAGT TGTAACCTCGT | 5100 |
| AAGTCAAATG ATGTCAATAG TCCACAATAT CGCGAACGTA CAGTAACGCT CAACGGACGT | 5160 |
| GTCAATCAAG ATGGTTCAGC TTAAGTGACT CCTTGGAAC TGGATGCAAA TGGTAAGAAA | 5220 |
| CTTTCTACTG ATAAGGAAAA GATGTACTAC TTCAATACGC AGGCCGGTGC AACAACTTGG | 5280 |

975

| | |
|--|------|
| ACCCTTCCAA GCGATTGGGC AAAGAGCAAG GTTTACCTTT ACAAGCTAAC TGACCAAGGT | 5340 |
| AAGACAGAAG AGCAAGAAGT AACTGTAAAA GATGGTAAAA TTACCCTAGA TCTTCTAGCA | 5400 |
| AATCAACCAT ACGTTCTCTA TCGTTCGAAA CAACTAATC CTGAAATGTC ATGGAGTGAA | 5460 |
| GGCATGCACA TCTATGACCA AGGATTTAAT AGCGGTACCT TGAAACATTG GACCATTTCA | 5520 |
| GGCGATGCTT CTAAGGCAGA AATTGTCAAG TCTCAAGGGG CAAACGATAT GCTTCGTATT | 5580 |
| CAAGGAAACA AAGAAAAAGT TAGTCTCACT CAGAAATTAA CTGGCTTGAA ACCAAATACC | 5640 |
| AAGTATGCCG TTTATGTTGG TGTAGATAAC CGTAGTAATG CCAAGGCAAG TATCACTGTG | 5700 |
| AATACTGGTG AAAAAGAAGT GACTACTTAT ACCAATAAGT CTCTCGCGCT CAACTATGTT | 5760 |
| AAGGCCTACG CCCACAATAC ACGTCGTGAC AATGCTACAG TTGACGATAC AAGTTACTTC | 5820 |
| CAAAACATGT ACGCCTTCTT TACAACCTGGA GCGGACGTCT CAAATGTTAC TCTGACATTG | 5880 |
| AGTCGTGAAG CTGGTGATCA AGCAACTTAC TTTGATGAAA TTCGTACCTT TGAAAACAAT | 5940 |
| TCAAGCATGT ACGGAGACAA GCATGATACA GGTAAAGCA CCTTCAAGCA AGACTTTGAA | 6000 |
| AATGTTGCTC AGGGTATCTT CCCATTTGTA GTGGGTGGTG TCGAAGGTGT TGAAGATAAC | 6060 |
| CGCACTCACT TGTCTGAAAA ACACAATCCA TATACACAAC GTGGTTGGAA TGGTAAGAAA | 6120 |
| GTCGATGATG TTATCGAAGG AAATTGGTCA CTCAAGACAA ATGGACTAGT GAGCCGTCGT | 6180 |
| AACTTGGTTT ACCAAACCAT CCCACAAAAC TTCCGTTTTG AAGCAGGTAA GACCTACCGT | 6240 |
| GTAACCTTTG AATACGAAGC AGGATCAGAC AATACCTATG CTTTGTAGT CGGTAAGGGA | 6300 |
| GAATCCAGT CAGGTCGTCG TGGTACTCAA GCAAGCAACT TGGAAATGCA TGAATTGCCA | 6360 |
| AATACTTGA CAGATTCTAA GAAAGCCAAG AAGGCAACCT TCCTTGAGAC AGGTGCAGAA | 6420 |
| ACAGGCGATA CTTGGGTAGG TATCTACTCA ACTGGAAATG CAAGTAATAC TCGTGGTGAT | 6480 |
| TCTGGTGGAA ATGCCAACTT CCGTGGTTAT AACGACTTCA TGATGGATAA TCTTCAAATC | 6540 |
| GAAGAAATTA CCCTAACAGG TAAGATGTTG ACAGAAAATG CTCTGAAGAA CTACTTGCCA | 6600 |
| ACGGTTGCCA TGACTAACTA CACCAAAGAG TCTATGGATG CTTTGAAAGA GCGGTCCTT | 6660 |
| AACCTCAGTC AGGCCGATGA TGATATCAGT GTGGAAGAAG CGCGTGCAGA GATTGCCAAG | 6720 |
| ATTGAAGCTT TGAAGAATGC TTTGGTTCAG AAGAAGACGG CTTTGGTAGC AGATGACTTT | 6780 |
| GCAAGCTTA CAGCTCCTGC TCAGGCTCAA GAAGTCTTG CAAATGCCTT TGATGGCAAT | 6840 |
| GTGTCTAGTC TATGGCATAC ATCTTGAAT GGTGGAGATG TAGGCAAGCC TGCAACTATG | 6900 |
| GTCTTGAAAG AACCAACTGA AATCACAGGA CTTGCTATG TTCCGCTGG ATCAGGTTCA | 6960 |
| AATGGTAACT TGCAGATGT GAACTTGTT GTGACAGATG AGTCTGGCAA GGAGCATACC | 7020 |

976

| | |
|--|------|
| TTTACTGCAA CTGATTGGCC AAATAACAAC AAACCAAAAG ATATTGACTT TGGTAAGACA | 7080 |
| ATCAAGGCTA AGAAAATTGT CCTTACTGGT ACCAAGACAT ACGGAGATGG TGGAGATAAA | 7140 |
| TACCAATCTG CAGCGGAAC TATCTTTACT CGTCCACAGG TAGCAGAAAC ACCTCTTGAC | 7200 |
| TTGT CAGGCT ATGAAGCAGC TTTGGTTAAG GCTCAGAAAT TAACAGACAA AGACAATCAA | 7260 |
| GAGGAAGTAG CTAGCGTTCA GGCAAGCATG AAATATGCGA CGGATAACCA TCTCTTGACG | 7320 |
| GAAAGAATGG TGAATACTT TGCAAGATTAT CTCAACCAAT TAAAAGATTC TGCTACGAAA | 7380 |
| CCAGATGCTC CAACTGTAGA GAAACCTGAG TTAAACTTGA GATCTTTAGC TTCCGAGCAA | 7440 |
| GGTAAGACGC CAGATTATAA GCAAGAAATA GCTAGACCAG AAACACCTGA ACAAATCTTG | 7500 |
| CCAGCAACAG GTGAGAGTCA ATCTGACACA GCCCTCATCC TAGCAAGTGT TAGTCTAGCC | 7560 |
| CTATCTGCTC TCTTTGTAGT AAAAACGAAG AAAGACTAGT ATTTAGTAAA ACCTCTTAAC | 7620 |
| AAGATTACGG AAGCAGTCTC TATCTTTTCC AATGAGGTTT ATAGTACAGA AAAAGCCTGA | 7680 |
| GAAGATGTCT TCTCAGGCTT TTGTTAAGCA CATAAATACA ATAGTGCTAT GACAAAATCA | 7740 |
| CCCAGAAAAA TCTGGGTGAT AAATGTTATG GTTGTGCTGG TTGAGGATTC TGATTTTGT | 7800 |
| GATCAGGGGT TGTATTTGAT TGTGCGTAT TATTGTTAGG ATTGGTAGTC GTACTATTAT | 7860 |
| TTGTGCTTGG AGTGGTTGAG CTAGACTGTG AAGTTGAACT ATCTGATGAT GAGCTTGAAC | 7920 |
| TTTCAGTTGA TGGGGGTGTG TGTGGAGCAG GTGAGTTCCA CGTAGAACGA GCACCATTTT | 7980 |
| TAAATACGAA TTCTCCATTT CTGTAGAGCC CCTCTGGTAT ATTCCAATCT TCTGGATTGC | 8040 |
| TTCTTCAGCA CAGGTAGGTC ATCATAGAGC GGTAAACTTT GGCAGCGACC GTAAGGCCAT | 8100 |
| TGCCTACAAG TGGTGTGAGA CGGTTAGAAT AGCCTGTCCA TACAGCCATT GAATATTTAC | 8160 |
| GCGTATAGCC AGCAAATAGT TCATCAGGTG CTACAAATTG AGAGGTCTTG ATGTGGTTTT | 8220 |
| CAATTTCTCT GTCTGTATAG TTAGAGGTTT CTGTTTTACC AGCCTGAGGG AGCCAAGCAA | 8280 |
| GATAGGCATT TCGTCCAGTT CCATAAGTCA AGACTGTTTT CATCATGTCG GTCATCATAT | 8340 |
| AGGCTGTCTG TTCTTTCATG GCACGAGTTC CGACATTAGA GAACTCTTTT TCACTCCCAT | 8400 |
| CACTAAAGAC GACTTTATGG ATATACATTG GTTTATAGTA AGTTCCACCA TTTGCAAAGG | 8460 |
| CAGCGTAAGC AGCAGCCATC TTTTCACTAC TTGCTCCATA TTTTTTGTCT GATTCGGTTG | 8520 |
| TGTTACTTGA AATGGCATTG GAGTAGTGAA TACTTGGGTA GTCGATTCTT AGACCATTTA | 8580 |
| GGAAAGTCTT GGC GCGGTTG AGTCCGACCT TGTTTAGAGT TTCCACGGCT GGGACGTTTC | 8640 |
| GCGATTGTTG CAGGGCGTAT TGCAAGGTGA TGTTGCCAAA GTAGCCCCTA TCCCAGTTAT | 8700 |
| AAACAGGAGT ATTTGTCCCA GGGTAGTTAT AGGGCTCATC GTGAACGATA GTAGCAGTTG | 8760 |
| AATCGTAGAC ACCGTACTCC AAGCAGGAG CATAGTCTGT GATCGGTTTC ATAGTTGATC | 8820 |

977

| | |
|---|-------|
| CCCAGTCGCG GTTTGTTTCT ACTGCTTGGT TAATTCCGAA GGAAACATTA CTTGACTGAT | 8880 |
| GGCGTGCTCC TAGCTGGGCA ATGACTTTAC CGTTAGAAAC ATCAACAATG GTAGAAGCGA | 8940 |
| CTTGCAATTC ATCGTCTGGA TAGGCAACGT ATTCGTCTGT ATTGTAAATA TCCCACAGAT | 9000 |
| GTTTTGTAGC TTCTTGGTCT ACATTGTGT AGACATCCAT CCCAGTTGTG AGTAGGTTAT | 9060 |
| AGCCTGTTTC TTCTTCAACT TGATTGATGA CTTCTTGAG GTAATTATCC ATGTAAGCAG | 9120 |
| GGTAATTACT TGCTGATTG AGACTTTGTA GTCCATCAGT AATTGGTGTA TTGACTGCTT | 9180 |
| TCTCATACTG TTCAGCAGAG ATGTAGCCTT GATTTTTCAT TTCAGATAAG ACCAAGTTTC | 9240 |
| GGCGGTCTTG GGCTGCTTCT GGATGTGAAT AGGGGTGATA TTGGTTTGGT GCCTGAGGCA | 9300 |
| TTCCAGCCAG CAAGGCTAAC TGAGGTAAAC TTAAATTATT GAGGTCTTTA CCATAGTAGT | 9360 |
| TTTGAGCTGC TGCTGCATT CCATAGTTCC CATTAGACAT GTAGACCTTA TTTATATAGT | 9420 |
| AGGTCAAGAT TTCTTGCTTG GTTGCTTTT GTTCTAACTG AATCGCTAAC CAAGCTTCCT | 9480 |
| GAGCCTTACG AGAAATAGTC TGGTCGGAAG TCGAAGTTGA AAAGTAAGTC AACTTAATCA | 9540 |
| ACTGTTGGGT GAGAGTTGAT CCACCTTGA GGAATTGCT TTGCAGATTG CGCAAGAAAG | 9600 |
| CTCCCAGGAT ACGGATGGTA TCAATCCCC TGTGGTCGAA GAAGCGATGG TCTTCGATAG | 9660 |
| AAACGATTGC CTTAACCAA TCTGTGGGAA TATCATTAGC TTGGGCATTG ACGCGGCGTT | 9720 |
| CAGAACCCAA GTCAGCAATG AGTTGATTTT TATTGTCGTA GATTTTACTA GAAGTTGTTG | 9780 |
| CAACTAGTTT ACTCTCGGAT AGGCTAGGAG CCTTGCTAAC GTAGTAGAAA AAACTCCTC | 9840 |
| CGCCTAAGAC AATGGCTGCG ATAACCAAGC TTAAGAAGCT AATGCTCAGA TACTTGATTA | 9900 |
| GGCGCAGAAT CGTTGGTTTG TTCATCTGT TTTACCACCT AATAAATGTT CTTTGATAAC | 9960 |
| ATTGAGATAA GGAATTGAG GGAAGGCACC AGCCTTGATT TCATATCCAT ATTCTCGAAT | 10020 |
| ATATTCAAGT GGCATTGATT TTTGTCCCTT ATCTTGATGA TAGAAGCGAA TCAAATCGAA | 10080 |
| TGCCGGCAAT AAGTAGGTTT CTTGCTGAGA AGAAAAGTGA AGAAGGACAA AGCAGATTCC | 10140 |
| TTGTTGGGCA AGGACTTGTT CCATATGCTG AATCTGATGT GGATGAAAAT TTTTCATCGG | 10200 |
| AATCGCACGT TTTTGTTTTG TTTCTTGAC TTCAAAGTCG ATGTAATATC CATTATAAAC | 10260 |
| GCCAGAATAG TCCGTCGTTG AAGCTTGTCG AAAATAGGCT TCAACAATCT TGGCAGGACT | 10320 |
| TCGTTGTGGA TAGTCCACTT GTACGATTTG AATAGGAGTT GGTTCCTTAT GTATAACAGC | 10380 |
| CAAGCCCTGA GACAAATAGT AGTCGTTGGT AGCATTGATC ATCTTTTCAA AGGGTACCGA | 10440 |
| GCTCGAATTC GTAATCATGT CATAGCTGTT TCCTGTGTGA AATTGTTATC CGCTCACAAT | 10500 |
| TCCACACAAC ATACGAGCCG GAAGCATAAA GTGTAAAGCC TGGGGTGCCT AATGAGTGAG | 10560 |

978

| | |
|---|-------|
| CTAACTCACA TTAATTGCGT TGGCTCACT GCCCGCTTTC CAGTCGGGAA ACCTGTCGTG | 10620 |
| CCAGCTGCAT TAATGAATCG GCCAACGCGC GGGGAGAGGC GGTTCGCGTA TTGGGCGCTC | 10680 |
| TTCCGCTTCC TCGCTCACTG ACTCGCTGCG C | 10711 |

(2) INFORMATION FOR SEQ ID NO: 146:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11887 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 146:

| | |
|--|------|
| TACATTCATT CCATCGGCTA CTCCATAATA CTTAGATAAA ACCATAGCTG AAGTCGAATA | 60 |
| CGGATACTGT AAAGTATTAT CAATTTTAAT CAAATCATCA TTACCGATAA TACTTCTGAT | 120 |
| TGCTTTTGGT AGTATGAACC ATACGTTGGT GAAATCTCAG ATAATGAAGA ATCATTAGAC | 180 |
| TCTGGACCTT TTTCTAGTGT CTCACCTACC TCATATTCTT CACCCTTACT AGAAATAACA | 240 |
| CTCAAAGCAG ATACTGTCTGA TAACTGGCTA GCCAATAAAG TACTCGCAAT AATTGAAATA | 300 |
| CCCAATTTTT TATAAACAGT TTTCTTCATT ATGTATCCTT CCTAATGTAA TTATAGCGTA | 360 |
| CTATTCTAAA TTTCTTAATC TACTATAGAA TCAAGAAATC TACCACCTTC TTTAAATACC | 420 |
| CTCCATTATC ACATAAACAG GTAAACTTTT CAATTAATGA CTGCGCTTTT CAATCACGCT | 480 |
| AGAGGTACTT GCTTGCTTCT TTGATACTAA GTTCAGCCAT TCTTTCCTTG TTTTCTCAA | 540 |
| TAAAGCATGT TACCCAAAGT GGATTCGTTT TGGAGTAGTC TCGCAGAGTC CAGCCAATGG | 600 |
| CTTTATTGAT AAAAAATTCT GTTTGGTTCA AGTTATGAAG GAGAATCTTT TCCATTAATT | 660 |
| GAGTATTGGT CTTCTCTTTT CTTAACAAC TGGTGGTCAAT AGCGACACGT CTCAGCCAGA | 720 |
| TATTATCTGA TAGGCTCCAT TTTATACTCA ATGAAAATCA AAGAGCAAAC TAGGAAGCTA | 780 |
| GCCGCACTG CTCAAAACAC TGTTTGGAGG TTGCAGATAG AGCTGACGTG GTTTGAAGAG | 840 |
| ATTTTCGAAG AGTATTAAGA TTATTTCTTC TAGTTCAGGG TGTTCATACA CCAAACCCC | 900 |
| TACTACTCGA TCTAGGATAT CTACCGTGTC CCACAAGGAT TTTGTCACGA CTAACCTGCTC | 960 |
| TAGCTTAGGC AAATCGGTTT CCTTTAGATA AGACTGCATT GCTTTCAAAT AGTTAGCAGC | 1020 |
| CACATATTGG TATTTTCTAG GATCCTTTTC CCAGCAAGTG TCTGCAAAAT CCCAATCGAT | 1080 |
| AATCTTTGTT TTTTTCGCTT CTGGAAAATA TTTTATAGAG TTTATTTCTT TCAGGCACCG | 1140 |
| CAATACCTAG AAAAGAAAAT TGATGGCGCA TATAGGCTTC CATGGACCTT GCTTTTCTAG | 1200 |
| AGTCTTTTGC TGCTTCTAGC TCCTCAAGTA AATCTGCTAA ACTCATCTAA AACTCCTCTT | 1260 |

979

| | |
|--|------|
| GCCCCACCAA ATGGTGCTGA AAGGCATAGA CAGCCGCCTG GGTACGATCG CTGACTTCAA | 1320 |
| GTTTGGCAAG AATATTGGAC ACGTGGGTCT TGACCGTCTT GAGAGAGATA AAGAGGTCAT | 1380 |
| CTGCGATGCG CTGATTTTCG TAGCCCTTGG CGATGAGTTG GAGAACATCT CGCTCACGCG | 1440 |
| CAGTCAATTC TTCATGAAGT TCCATATGAT TGCGGTGGTA TTCAACCTTC TTGCTAACCT | 1500 |
| CTTGCTCAAT GGCCAGCTCG CCAGCAGCTA CCTTACTGAC GGCATGAAGC AATTCATCTG | 1560 |
| CACTAGAAGT CTTGAGCATA TAGCCTTTGG CACCAGCATC TAAGACTGGC ATGATTTTTT | 1620 |
| CATTGTCCAA ATAAGAGGTC ACAATCAAAA TCTTGGCTTC AGGCCATTCT TTAAGGATTG | 1680 |
| CTAAGGTGCG GTCAATCCCA TTCATCTCAG GCATGACAAT ATCCATGACA ATGACATCTG | 1740 |
| GACGCAGTTC CAAGGCCAAG TCAATCCCTT GAGACCCGTT GGACGCCTCA CCCACAACCT | 1800 |
| CTACATCGTC TTGGAGGTCA AAGTAGCTTT TCAAGCCCAA TCGGACCATT TCATGGTCAT | 1860 |
| CTACTAGTAA AATTTTCATC TTTACTCCTT TATCATTCCT TATCTAACAG GGAATACGG | 1920 |
| ATATCAACCG CCAGCCCTTG CTTGGGAGCT GTCAAGAGTT GAACTGTTCC AGCCATATCT | 1980 |
| TCAACCCGCT CCTTGATATT TCGCAGTCCA TAACTCAAGT CGTCTAAGCT CCCTAACTGG | 2040 |
| AAACCAATCC CATGTGCCAC CACCTTCAGT TGCAATTCAA CATCTGTCTG ATAGAGGTAG | 2100 |
| ACATCTAGGC AAGATGCCTG GGCATGGCGG AGGGTATTGC TAATCAACTC TTGCAGGATA | 2160 |
| CGGAAGATAT GCTCCTCGAT TTTCTTAGGC AATTTCTGTC TATTCTGCTT GAGACTAACC | 2220 |
| CTAAGATCAC TCTGTCTCTC AAGCTCTTTT AAAAGAATTT GAATCCCTTC TATCAAGCTC | 2280 |
| TTCTGCTCCA GTTCAACTGG TCGCAAATGC AAGAGCAAAA CCCGCAAATC CTTCTGGGCT | 2340 |
| GTTTCTAAAA TAGCTGTGAC ACTCTGCAAC TGGGTCTGCA TCTTTTCTCT ATCCAATTTT | 2400 |
| AAAGCCTGCT GACTGATACC CGATAAAATC ATGTGGGCCG CAAACAATC CTGACTGACT | 2460 |
| GTATCGTGCA AATCCCGAGC AATTCGCTTC CGTTCCTTCT CGATGATTTC CTCTTCCTGA | 2520 |
| GCAAGGCTCT GATTTTCAGC TTTTGAAGA GCCTCTGTCA AAAGGTTAAG TTTACCTGAT | 2580 |
| AAGGACTTGA AACTGGCATC CAAATCTGGA TCTGCAACCT GAACCACTTC TTGCCCTGCT | 2640 |
| AATAAACGCT TGAGATTAGC CTGCATTTTT CTTAGAGAAA GCTCTTCGAT CCCTCGCCAA | 2700 |
| AACAGGGCTA AGAGACAGGT CATGGACATG CTGAAAACCA ACAATAAAAA GACAAATTTT | 2760 |
| TCTGTTTTTT CGACATCGTG CAAAAAGATA GACCAGTCAA AATCAAGTAT TTCCAGCAAG | 2820 |
| CTGTGGGAGA AAAAAAGAC AAATAGGAAG GAGGTGAGAG CAATAATGAC ATAGGCTTGT | 2880 |
| TTTTTCATCC TCTAACCACC TCCACATCAC CAATCATAGT GGTCAAGAAA ATCTTGACAC | 2940 |
| TCTTGTTACT CTTGAGATAG TCTTTTGTTT CTTGATGATA GTGTTCAATTG CGGAGGGCTC | 3000 |

980

| | |
|--|------|
| GCTTGGGCTG GTTGAAAAA ATCAAATCCC CATAGAGACA GTTAACGCTG AGACTGACTT | 3060 |
| CCACATCTAC AGGTACGATG ATTTTGGTCG TTCCTACCAT CTTTCTGAGG ATAATGACAT | 3120 |
| TGTCATGATT GGTAAAGATG ACCCTCTCCA GATGAATAGT GTCCTTGCCC ATGAAGCGAA | 3180 |
| AGAGATTGAT ATCATCGAAT TGGCAAGTCT GGTAGCTTGA AAAATGATGA AGATTTCCAA | 3240 |
| ACCAACGATT TTTCTCCTTC TTAACCGTCA CGACCTCTTC AAAAACCAAA TTGGTCTGCT | 3300 |
| CTTTTTCCTG GTTCATCATC GGGTAAAGAA GAAAGAGGCT ATAGATAACC GCAACAAAAA | 3360 |
| TAGCTAGAAT CACAAAAGGA TTGAGCATAA CGATGAAAAA GAAGAGAATG GTTGCCGCTA | 3420 |
| CTAAAAGAAG ATTATTTCCC TCTTTACCAG TGTTAGTAGCG AATCAAAAGC AAAAAGAGGA | 3480 |
| ATAGTATCAG CAGAAAACGC GAAAAATGCT CTGATACCAT CAAATCAGA GCTCCTGTCA | 3540 |
| GAAGACAGGC TTCGATAAAT AAAAAGATTT TAAATTTTCT CATAGGTTCA TCCTCTCCCT | 3600 |
| TCTATTTTAT CACAATTCAA AAAAGTCACC TCAGTCTGAG GATGGAAAAA AGGCGCTGGT | 3660 |
| TACGCCCTTT TCATCTGATC CTTTGCTTCT TTTAATTTTC CATAAAGAAG ATAGTCTACT | 3720 |
| TTTGTAGAT CTGCTATGGT GGCACAGTTA AGGGAACACA TAATCAAGCG TAGATCTGCT | 3780 |
| TTCCAGCCTT GGACAATGCC AATCACTTCT TCAACTGTGT AGGTTTCAAC CAATCCAGA | 3840 |
| ACGGTTCGTG ACAATCCAC AGCCTTAGCA CCAAAAACCA AGCACTTAAT CATATCCAGC | 3900 |
| GGATTCCGAA CCCCTCCACT AACCAAGAGT TCGACCTTAT CTTTCCATTC TTGGGCATG | 3960 |
| AGAAGGGCCT GCATGGTAGA CTGACCCCAT TGATTGAGGT AATCACGCTG GCCACTACGA | 4020 |
| CGGTTTTCGA TATAGGCAAA GCTGGTGCCA CCACGACCG ATAGGTCCAC TGTACGAACA | 4080 |
| CCGAATTCAT AGGCTCTTTC GATTGTCTTG GCATCCATTC CAAAGCCAC TTCCTTGAGG | 4140 |
| ACAATAGGAA CGGGAATTG CTGCTATAA TCTGCTAGAT GCGATTGCCA GCTTCTAAAC | 4200 |
| TTCTTTCTC CCTCGGGCAT GAGTAATTCC TGCATGACAT TGACATGCAC TTGCAATAGA | 4260 |
| ACAGGATTCA TCTCTCTAC AGTCTGAAGT CCTAACTCGA CAGGCTTGTC CAATCCAATA | 4320 |
| TTGGTTCCAA GGAGGAGATT GGGATGACTA GACTTGACAG AAAAAGAATC ATCCGTTGGA | 4380 |
| TTTTTGAGGG CTGCGCTATA AGAACCGTT ACAAATAAAA TACCACAGGA TTCCGCCACC | 4440 |
| TGAGCCAGCT TTTGATTGAT TTCTCTTCCC TTATTACTTC CACCAGTCAT GGCATTGATA | 4500 |
| TAAAAAGGAA AGTCCCACTT TCGACCAGCA AACTCTGTCG AAAGATCGAT TTCATCCAGA | 4560 |
| TTGTAAAGAG GCAAGGAAGA ATGAATCAGC TCCACCTCAT CAAAGCTATT ATAGGAACTT | 4620 |
| TTCTGCTCAA GGCATAGAG GATATGCTCG TCCTTACGAT TTGTCGTCAT GTCCTATCCT | 4680 |
| TTCTTGATAT AAGAGCTCAA TCCCAGATC GGCCCAACGA TTTTPTAAGG TTTTGGTTGA | 4740 |
| TTGCGCATCA AACTCAGGG CGATGCCACA GTCACCACA CCAGCACCAC TACTCTTGGC | 4800 |

981

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| AACGGTCTGC | AAATCTTGAC | TGGCTTCTTT | CAACTGTCTA | AGCAAAGGCG | TGTAAATATC | 4860 |
| TGTACTCAAG | CCTTCTAAAA | GCTTGCTGGC | TACTTCTACT | TGATCGATAA | TCTTTTCTGA | 4920 |
| TTTCCCTGT | TCCAAGGCTT | CTACCAGAGA | AGTCACCGTT | TCTTTTGAGG | AAGTTAAAAA | 4980 |
| ATTTTGATTG | ATATTTTGCT | TGATTTGCTG | GACCATGTGA | CTCGATACAG | CCACTTCCTT | 5040 |
| GGTCCATCCC | ACTAAGAAAT | CACATTCTAA | AGTTGGTTTC | ACTTGTGAAA | TTGAAAAGCC | 5100 |
| CCAATCACGC | TCCAGAACTG | TCGCCAAGTT | TTCTTCTTCT | AACCAAGCAG | CCACCTTCTG | 5160 |
| GCGATCAAAT | GACTGGTAGA | GAACCAAATC | CTCTGCCACA | ATACAGGCAA | GGTCGCCCAT | 5220 |
| GGAACCATTG | TCTCCTCGCT | TAAGCAAGAC | AGCGCTAGTC | AGCTTGAACA | AGAGCTCCTG | 5280 |
| ATCAACAGAA | ACATCATACA | GAGCCAGTAA | AGCCTTGACA | ACCAAGACAA | CGACGCTGCC | 5340 |
| ACTAGAACCT | AGACCAAAT | TTTTCCCTTC | TCGTTCCATT | TTGCCACAGA | TTTCTAGAGA | 5400 |
| AAAAGGTCTT | AAATTCTGAC | CACGAACAGC | GAGGAAGTCT | CCCATCAAAG | CAATCGTTTC | 5460 |
| TTGAATCAAG | CTATAGTCAG | GATTAGGCCT | TAAGTCCACT | GCGAAATCAA | ACATATCTGA | 5520 |
| ATAGATACGG | TAGCTGTCAG | AAAAAGCAAT | CTCAGCCCTC | ATATAGATGG | GAATATCCTT | 5580 |
| TATCAAAGCT | AAGTCCCCTG | GCTCTAAAAT | AGCATATTCA | CCTGCCCAAT | AGAGTTTTTC | 5640 |
| GCAAGTTTTA | ACAGCAATCA | TCTTGACTCA | AATCCTTTGT | TTTTGACACA | ATCAAGCGAT | 5700 |
| AACGATGACC | GAAAATTTCT | GATAAATGCT | CCAAGTCTTT | CTCCTGACAG | AAGACCTTAA | 5760 |
| CATTGGGACC | AGCATCCATG | GTAAAGTAGC | AGGCCTCTCC | TTTCTCACGA | AGCTGGCGAA | 5820 |
| CAAAGGCCAT | AGCCTCATAA | GAGGCATCCG | TCAGATAAGA | AAAGGCTGGA | CTAGCAGTCT | 5880 |
| TTGTCGTAGC | ATGCATAGCC | AGGGCATTTT | TCTCCGTAA | TTCTCCAATC | TTGGCAAAAT | 5940 |
| CATTTTCCTT | GAGATAAATC | AGCATATCCT | GATAGTCCTT | CTCAGACTGA | CGAACCCAGT | 6000 |
| CGTCGAAAGT | CGTCGAGGTT | TCCACACAAA | GTTTCATCCC | GTCACGGCTA | GAGATTGGTT | 6060 |
| TTTTCTTGTC | CTCTAGCACC | AACATAATCA | TAGCTAGTTT | CAAGTCTGTC | TCTACAGGGT | 6120 |
| AAATTTCTCC | ACTATCCTTA | TCCCAGGCTC | CTAGTGGTCC | ATAAAAACTC | CGAGAAGAAG | 6180 |
| AACCTGAGGC | AAATTTGGCT | TCCTGTGCCA | ACTGACTTCT | ATCCAATCCA | AGCTTGAAAT | 6240 |
| AAGCATTACA | AGCCTTGACC | AGGGCGGACA | AACCACTAGA | ACTTGAGGAC | AGACCCGCTG | 6300 |
| CCGTAGGCAT | ATTGTTTTGA | GTATCGATAC | GGACAAAGCC | CTCACCAGCT | GGACGATAAC | 6360 |
| GGTCAATAAT | CTTACTCATC | TTGGCATGCT | CGACCTCATT | TTGTAGCTGA | CCATTGATGT | 6420 |
| AAAATTCGTC | AGCTGTTACA | TTGGCTGGTA | AAGGCGACAA | GGTCGTCTCT | GTATACATAT | 6480 |
| TTTCCAAAGT | TAGAGAAATA | CTGCTAGTAG | CAGGCACCAT | CTCTTTTTCT | TTTTTCTTTC | 6540 |

982

| | |
|--|------|
| CCCAATATTT GATAATAGCA ATATTTGCGT AGGAACGTAC TGTACAGGC TCTCTATCCA | 6600 |
| TGTCTGAACA GCTCCTTTCT CTTCTAATCT TTCTGCTAGT TCTTGTGCGT GTGTCAAATT | 6660 |
| GGTTACCAAG GCTATGATAC AACCTCCTAG CCCACCACCG CTCATCTTGG CACCCAGAGC | 6720 |
| ACCATGGCTA AGAGTCGTTT CAACCAAAAA GTCTGCCTCA GGGCTACTGA CTCCAATTTC | 6780 |
| TTTTAAATGT AAATGCGCTT GACTGAGGAT TTGTCCCAGT CCTTCAGCAT CTTTTTGTGA | 6840 |
| AATCGCAACT TCTGCTTGCT GGGTTAATTC TCCCAAGGCA TGCAAAAACG GTAGGGCATC | 6900 |
| CTTGCCCTTA TTTTGAACCA CTTGGATGGC TTCACGAGTA TGACCATAAA CACCCGTATC | 6960 |
| GGCAATCACC AAATAGGCGG ATAAATCCAT CTCAGTTTCT GTAAATCCTA CGTTCCTGAT | 7020 |
| AAAGCGAATA GGTGGTCAC TAAGACAGGT CTTAGCATCC AAACCACTAG GATTCATATG | 7080 |
| GGCAATCATT TCAGCTCGAT TGACCAAGAT TTCTAGTACA TCATGAGGCA GATCAGCCTG | 7140 |
| ATAGTAGTCA AATACTGCAC GAATGGCCGC TATGCTGATA GCCGCTGACG AACCCATCCC | 7200 |
| CCGTTTCTCA GGGATAGCCG AGTCAATCTC ACAACGAATG CAGGCTTCTG TGATATTCAA | 7260 |
| ATACTCCAGT GAGGCATAAA CCGCCATGGA CAAGGTATCC TCCTCATAAA GGCGCCAAGG | 7320 |
| ACTCTCTGCA GGAACACTCT TACAGGTCAC CTCCACCTCC AAAAGAGGCA GGGAAATGGC | 7380 |
| AGGATAACCG TAAACGACCG CATGTTCCCC TATTAATAAT ATCTTACTAT GTGCCTGACC | 7440 |
| GACACCAACT TTTTTGTCA TTTTTCCTT TTAGTAGACG AAAAAACGTC TTATTTTTCA | 7500 |
| TACAAGTATT AATCTTTCC TATCTATTTT ATTATATTTT CACAAAAAAA GCGATTGTTT | 7560 |
| CCATTCACAA TCGCTTCTTT CATATTGAA CCCATTGCGC ATTATAGTTG ACAGAATAGC | 7620 |
| CATCTACGGT CGTATTCACT GCCAAGGCAC CTGAGCGCTA TAAGCGTAGT ACCATCTGCC | 7680 |
| ATTGACCTGG AACCAACCTG TCGTCATAGA ACGACGAAAG AAACCTCCATA CCATTAAGTA | 7740 |
| AAGAGGAAAG TCGTGAGGGA GCATGCGCCA TTGACAACCT GTTTTAGTGA CGTACAAAGT | 7800 |
| CTCATTAACA AGTACTCGTT TCGGCCATTT ATAGGTGCGG TGTTTGAGGA AATAGGGTTC | 7860 |
| AATCTTCGCC CATTCCTGAT CGTTTAAATC AGTATCATAT GCTTTGCGTA TCATAACTCT | 7920 |
| AGCTTAACAT TTTTTGTGA ATACAGGTTT TAAATAATCG ACCACGAAAA TTTCTTAAGT | 7980 |
| GGAAAACGCC TTATGAAGTA TGCTACGGGA AAGTTATGCA CTTAATTGA CAATTCAAGA | 8040 |
| TGTAAAAATA TATACTATAG TAGATTGAAA CTAGAATAGT ACACCTCTAC TTCTAAAAATA | 8100 |
| TTGTTAGAAA TCGATTTGAC TGTCTGATC GATTATCCT GTTATTATCT CATTTTACTA | 8160 |
| TAATATTTGA TAAGTTATCC TAAAAGTATT ATTATGTTGT TGTGTTATAG ATTGATTGAA | 8220 |
| TCTAACTAAA GGATCCTATT CAATTACTAG AACTATCACA TACTCAAGGT CAGCTCACAG | 8280 |
| ATGAGCAACT ATTTTGGTTA CAATGTCTAC TAAATTTAAG TCAACAAAT AATTTAGTCA | 8340 |

983

| | |
|--|-------|
| AAATTAAGGAA AATAGAGGAA CATAAATATG ATTACAAAAC AGAATGTAAT ACTGTTCTAC | 8400 |
| AATTTTACT AGATAAACT GTAAATTCTG AAGGAAGGAT CACTTCTTCA ACAGAATTTG | 8460 |
| GAAATTCGT AAGTAATTTA TCATTCCAAC ACGGAATAGC TGGACTACTG TTTCCTCTAA | 8520 |
| ATAAATTGTA CCCCCAGAA CTGGATTCTA AAATACTCTC TATCATCAAG AAGGCAGTGA | 8580 |
| CAATTAGAAC GACACACACA TATGAATATC AATACTCACT GCTATTTGGT GATGCAGGCT | 8640 |
| ATCTATGGTT ACTCCTACAT TTATTTCTA TCAGTAAAAA TCAATACTAT CTACAATTAG | 8700 |
| CAAACGTCAC CGCTAAAAA TTAATAGAGA ATTATGATAC TCTAGAGGAA ATAGACTTTG | 8760 |
| CATTGGGAAA ATCTGGTGTC CTATTATCAT TAATAAATA CTATCAATTT ACCAATGACA | 8820 |
| ATACTCTTAA AATTTTCATC CACAATAGTA TAGGGGAAAT TTATCATTAT TTCCTACAAA | 8880 |
| GAGATACAGC CAAAGAAAGC ATTTTAGACT ATAGCTTTGC TCATGGATAT TGTGGAATTG | 8940 |
| CATATGCTTT ATTTGCCTAT TCTAAAGTCT TAGAACCTTC TATGTTTAT AATGATCTCC | 9000 |
| ATACATTCCA TACTGAATTA AAAAAATTAT TAGAAAACT TACTTCTAAT ACTGAAAATT | 9060 |
| TAGGAAATTT ACAACTTTCT TGGTGCAAAG GAATTTCCGG AATAATCTTA TATCTTTGTA | 9120 |
| TGTACGATTG TGACGGAAAC AAAGATATTA TTAGTAAATA TCAAGAATTT GTTTTTAACC | 9180 |
| ATCATCTAAA AATGATGACA GGATATTGCC ACGGAATAAC TAGCTTACTA CAAACCACTG | 9240 |
| TCTACAATCA AAACAAATTA CTGATGAAAA AAATCCAACA GGTAATTTTA GCATGTTCTG | 9300 |
| AACGAGATGA TCACGGTTTA CTGATGTTTC AAGGAGATAG TGGTAAAGCA GATTTGTTTG | 9360 |
| ACTTCGGAAT AGGAAGCATG GGGTATATTG GTGCTATTA AATAATAAAT TCCCATTGCA | 9420 |
| TGTGCAGACA TAAGGAGAAA AGTATGAAAT TATTTTGGAC AAACAACATA TATAGACAGT | 9480 |
| TGCTGCTAAA CAGCTGTTTT TCATCATTCG GCGACAGTAT TTTCTACCTC GCCATTATCA | 9540 |
| ATTATGTGGC TCAGTACAAT TTCGCTCCGC TAGCGATTTT ACTGATTTCC ATTTTCAGAGA | 9600 |
| TGGTTCCTCT ACTATCGCAA CTCTTCTCG GGATTCTAGG AGATTTTCAA GAAAAAGAG | 9660 |
| TCAAACACGC ACTCTGGATT GCCAAAATCA AAATCCTGCT CTACGCTATT TTGACAGTAT | 9720 |
| TTCTCGTCTT GTCGCCCTTT TCATTAGTTT CAGTCATTAT GATTGTCATC ATCAACCTCA | 9780 |
| TCTCTGACAC CTTGAGCTAC CTGTCTGCCT ACATGATGAA CGCCCTCTAC ATCAGTGTAA | 9840 |
| TTAAGGACGA CCTGCATGAT GCCATGGGGT TCAGGCAGTC TCTGATGAGG GTTGTCCGTA | 9900 |
| TTGTGCGCAA TCTGGCTGGC GCATTCTTA TCAATGTTAT AAGTATTCAA ACTATTTCCC | 9960 |
| TTATCAACAC TCTGACTTTT GTCATTGCCT TTTTGGGCCT GTATGTTATT CGACATACCT | 10020 |
| TGTATGAGGT TGAAAAAGA ATTGAAATGT CACATACAGC ACTGAGTTTT AAGAAATATT | 10080 |

984

| | |
|--|-------|
| TTCAACATCT TAAACAGTCG CTGGCTGTGC TCCTGAGGTT AAAAGATACC GTCATACTAC | 10140 |
| TGTTTCTGAC GACCAGTATG ATTGCCATCT TGGATGTGTC CCCTCGGCTG ATTGCCCTCC | 10200 |
| GCTTCATCCA ACAGACACTA GCACAACGA GCATTGGGCA ACTCCTCGCC CTGCTCTCCA | 10260 |
| TCATCATGTC TTGTGGAGCT ATCCTTGGCA ATATGACCAG CAGTAATCTA TTTAAAAATA | 10320 |
| TCCGTTTCAC GCACCTCTTG GTTTTCTGTG AGATTTCCCT ATTGACTCTA ATAAC TAGTA | 10380 |
| TCCTTTGTCA AGCCTATATC GTAATTTTCA TGACCAGTTT CATCAGTTCT ACGATTATCG | 10440 |
| GCATTCTCAG CCCTCGCCTA CAAGCAGCTG TCTTTGCCA TATCCCCAGT GACAAGATGG | 10500 |
| GGACGGTGGG CTCTGCTCTG AGCACAGTGG ACATTCTCGC CCCGTCCCTG CTCTCCCTAT | 10560 |
| TAGCCCTATC CATAGCATCG GCGTTTCGG TGCAGTTAGC ATTGATATTT TTGTATCTTA | 10620 |
| TTTAAATTGC TCTTATCTTT TGTC AATGGT TAGTCAAGTT CAACACTCAT AACTAACGAA | 10680 |
| AAAGCATGTG TAGATTTTAC ATGCTTTTAA TCTCCCAAT CGTCAGGTCA AGTACAACAA | 10740 |
| AGTCACTTCT TTGATTAAGC GAGTGTCTA ATATAATTAT AAGCGCCCTG TCATTACCGA | 10800 |
| ACCCATTCGC CATTATAGTT GACAGAATAG CCATCTACGG TCGTATTCAC TGCCAAAGCA | 10860 |
| CCTGAGCTAT AAGCATAGTA CCAGTTGCCA TTGACCTGGA ACCAACCTGT CTTCATGTCT | 10920 |
| CCATTACCTG CATTTAGGTA GTACCAAGTT GAACCATCTT GATACCAACC AGTTGCCATA | 10980 |
| GCTCCTGATG AACGGAGATA GTACCATTG TTCCCAAGGT TTTGCCAACC TGTMTTCATA | 11040 |
| TCGCCATTTG GGTGGTCTAA ATAATACCAA GTGGTACCTT CCTGATACCA GCCAGTGGCC | 11100 |
| ATTGCTCCTG AGGAACGGAG GTAGTACCAC TTATTACCTA GATATTGCCA ACCTGTTTGC | 11160 |
| ATAATACCAG TTGTTGGATC TAGGTAGTAC CAAGTCGAAT CATCGTTTAT CCACCCGCA | 11220 |
| CGTCTTTCAC CACCAAGGTA GTTTTCTCCA TTAATTTCCG TCTTAGCTAG ATAATACCAG | 11280 |
| TTAGACTGAT CATAAAGCCA ACCTGTCTCT AAAGAATGAT TTTGATTAAA GTAATAGTTC | 11340 |
| GTATAATAAC GCTTCTCTTC TTTATCTTCT GAATCTTCAC GTTTTTCCTT GTACTTTCTT | 11400 |
| CCAACACTGT CTTTAGTTTT AATCTCTAAT GTTTTCCAAC CAACAACTC TTGTAGCACT | 11460 |
| CCATTTTAT CGAAGTAGTA CCACTCTGAC TTTGGAAAAC CTTCTAATCT GATACCATTT | 11520 |
| GGGTAAGGAC CAATTGTACT ACCTTTAGAT GGAACGGGA TATATTGCCA GCCGACAACC | 11580 |
| ATCTCTCCAG ATAGAGAATC AAAATAATAG TACTTACCAT CAATCACTCG CCAGTAGGTT | 11640 |
| TCTTTGAGGT CCCCCTTTTT GTAGTAGGTT CTTCCGTTTT CTTGGACAAA CTGCCATCCT | 11700 |
| TCAGAAATCAT CTGCAAATAC TGTACTGGTC CTAGCAAAAC CAAAGAAAAA TACTGTCACT | 11760 |
| CCAACCTGCA TAGTTTTTTT CAAAATTTTC ATCTATATAC CCTCCAATAT TAAATCCACT | 11820 |
| CACCAGATGA GGCGAAATTA TAACTTTTAC CATCGATAGT TTGGCTACCT GTAACCATG | 11880 |

985

CTCCAGG

11887

(2) INFORMATION FOR SEQ ID NO: 147:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11340 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 147:

| | |
|--|------|
| CCGGTATGTT CTGGAATACT ACCAATCTAA GCTGGCTGTG CCCTACAGTT TTACAACCCCT | 60 |
| GTACGAATAC CTTAAGGAAT ATGACCGATT TTTCAGCTGG GTTTTGGAGT CTGGTATTTC | 120 |
| AAACGCTGAT AAAATATCCG ATATTCTTT ATCAGTTTGT GAAAATATGT CTAAGAAAGA | 180 |
| CATGGAATCC TTTATCCTTT ATCTACGTGA ACGTCCCTTG CTGAATGCTA ATACAACAAA | 240 |
| ACAAGTGTT TCACAGACAA CTATCAATCG AACCTTATCA GCACTTTCTA GTCTTTACAA | 300 |
| GTATCTAACC GAGGAGGTTG AAAACGATCA GGGGGAACCT TATTTCTATC GTAATGTAAT | 360 |
| GAAAAAAGTT TCCACCAAGA AAAAGAAAGA AACCTTGCT GCCAGAGCTG AAAATATCAA | 420 |
| GCAAAAATC TTTCTAGTG ATGAAACAGA AGTTTCTTA ACTTATATCG ATCAAGAGCA | 480 |
| CCCACAACAG CTTTCAAATC GAGCTCTCTC ATCATTCAC AAAAATAAAG AACGAGATT | 540 |
| AGCCATTATT GCCCTTCTCT TGGCATCTGG TGTTCGCTTA TCTGAAGCTG TTAATCTAGA | 600 |
| TCTAAGAGAT CTCAATCTAA AAATGATGGT TATTGATGTT ACTCGAAAAG GTTGCAAACG | 660 |
| TGACTCAGTC AATGTCGCTG CTTTGTCTAA ACCTTATTTA GAGAATTATC TGGCCATTTCG | 720 |
| GAATCAACGC TATAAACCG AAAAACAGA TACAGCCCTT TTTTAACTC TCTACAGAGG | 780 |
| TGTTCTAAT CGTATCGATG CTTCTAGCGT TGAGAAAATG GTTGCTAAAT ACTCAGAGGA | 840 |
| TTTTAAAGTG CGGTAAACAC CCCATAAACT GCGCCATACA CTAGCAACTA GGCTCTATGA | 900 |
| TGCGACTAAA TCACAAGTTT TAGTCAGTCA CCAACTAGGA CATGCTAGCA CACAAGTCAC | 960 |
| TGACCTCTAT ACCCATATTG TTAGTGATGA ACAAAGAAT GCTCTGGATA GTTTATGATT | 1020 |
| TTACGTATTT TAAATTATGT AAATAAATAT CAAAAAAGA AGTTGGCCAA CTTCTTTTGT | 1080 |
| ATTTATCCAA CTACCGCTTC AGCGATTCTT TCACGGCTAA TACCAGCGAA GTAGCGTGTG | 1140 |
| ATATCAATGG TTTTATAGCG CTTAAGAACA TCTTCGCGTT CGTATTTTAC CCCACGAAGG | 1200 |
| ACATCTTCTA CTGCAGCAAC GTCTTCAATA CCAAAGAAGT CACCATAAAT CTTGATGTCT | 1260 |
| TGGATTTTGT ATTCAGTAAC GTTAGCAAAG ACTTCAACCT TACCACTAGT GAATTTGATT | 1320 |

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| 986 | | | | | | |
| CCACGACGGA | CGTTAAATTC | AGGTGATTTA | CCATAGTTCC | AGTCCCAAGT | TCCAAACTTA | 1380 |
| GTATCCTTGA | TGCGATTGAT | TTCGGCCAAT | TCTTCTTCTG | AAAAGACGTA | TTCAGTCATC | 1440 |
| TCTGGGTACT | CTTTTTTCAT | GTATTCCAAG | AGTAAATCAC | GGAATTTTTT | GACTGTGATT | 1500 |
| TTTTTTGGTA | ATTCATTGAT | AATATTGGTT | ACACGGGCAC | GGACGGATTT | CACACCTTTT | 1560 |
| GATTCAAATT | TATCTTTTGA | AACCTTAAGG | GCATTTGCGA | GGACTGACAA | ATCAACGTCA | 1620 |
| AAGAGCAAGC | AACCGTGGTG | CATGATACGG | CCGTGATAT | AGGCTTGGGC | ATTGCCACAG | 1680 |
| AACCTCTTAC | CATCAATCTC | AAGGTCATTA | CGACCTGTGA | ACTCAGCTTT | AACCCCAAGT | 1740 |
| TGAGCCAGGG | TATTGATAAC | CGGAGTTGAG | AAGCTCTTGA | AGTCAAATGC | CTTATTTTCA | 1800 |
| TCTTCTTTGG | AGATGATCGT | GTAGTTGAGG | TTATTTAAAT | CGTGGTAAAC | AGCTCCACCA | 1860 |
| CCACTAATAC | GGCGAACTAC | CTCAATACCA | TTTTCGCGAA | CATAATCACG | GTTGATTTCT | 1920 |
| TCGATAGTGT | TCTGGTGACG | ACCAACAATG | ATAGATGGCT | TGTTAATCCA | AAGTAGGAAG | 1980 |
| ATTTGATCCT | CATCCAAAAG | GTGTTTAAAG | GCGTATTCTT | CCAAGGCAAT | ATTAAAAGCA | 2040 |
| GTGTCATTTG | AATGATTGAT | AATGTATTTC | ATGATATCCC | TTTACTTTAT | ATGATAGAAA | 2100 |
| CTGGAAATAA | CCTTCCAGTC | TAATCTATCT | TCGTTTTATT | TTTTCTTAGG | TGAATGGATG | 2160 |
| GCCATTCTTA | GAACATCTGC | AAACGCTTCG | TACATCACTT | CAGAGTAAGT | TGGGTGCCCC | 2220 |
| TGGATGGTCT | TCAGCATTTT | CTCAACAGTG | ATTTCCATTT | CGATGATGCT | TGATGCTTCG | 2280 |
| TTTATTAATT | CTGCGGCTGC | AGGACCAATA | ATGTGTACAC | CAAGGATTTT | TCCGTATTTT | 2340 |
| TTATCAGCGA | TAACTTTTAC | GAAACCTTGA | GCTGCGTCAG | ATGCAATAGC | ACGACCGTTA | 2400 |
| GCAGCAAAGT | TAAACTTACC | GATGGCAACA | TCGTATTTCT | CACGGGCTTG | TTCTTCTGTC | 2460 |
| AAACCTACTG | CTGCTACTTC | AGGGAGAGTG | TAGATGGCTG | CAGGAGTCAA | ATTCAATTTG | 2520 |
| GCAACTGCAT | GATTTCTTTT | AAGGGCATTT | TCAGCGGAAA | CTTCACCCAT | GCGGAAAGCT | 2580 |
| GCGTGAGCCA | ACATCTTAGT | ACCGTTGATG | TCACCTGGTG | CATAAATGCC | TGGAACTGAA | 2640 |
| GTTTCCATGT | ATTCGTTGAC | CTTGATACAA | CCACGATCCA | ATTCAAACTC | AACCTCTCTA | 2700 |
| ATACCTTCAA | GGTCTGGCAT | ACGACCAATT | GAAAGAAGAG | CTTTGCTTGC | GATGATATCG | 2760 |
| TCTTTTCCTT | CAACCTTGAT | ACGAAGTTGA | CCATTTTCCT | CAATGATTTT | TTGCAGTTTA | 2820 |
| GTACCAGTCA | AGATGGTCAT | TCCTTTACGC | TCAAGAATCA | AGCGAAGGTT | CTTAGAAACT | 2880 |
| TCCACATCCA | TAGCTGGAAC | TATACGGTCC | ATCATTTTGA | TAACAGTCAC | TTTTGAACCA | 2940 |
| AATGTATGTA | AGGCCTGACC | GAGTTCGATA | CCGACAATC | CACCACCGAT | GATAACAAGG | 3000 |
| CTTTCTGGCA | CTTCGTTTAT | TTCAAGAATG | TCATCACTAG | TCATGACAAG | TGGAGATTCC | 3060 |
| ATACCAGGGA | CGTTGATCTT | GTTGACTTTT | GAACCACCAG | CAAGAATGAT | TTTCTTGGTT | 3120 |

987

| | |
|--|------|
| TCAAGCAATT CAGAACCATT TACCAAGACG TTCTTGTCTT TAGTGATTGT ACCAATTCCT | 3180 |
| TTATGAACAG TAACTCCGTA GCTACGAAGA AGTCCTGCAA CACCACCAAC AAGAGTATTA | 3240 |
| ACAACTTTAG ATTTAGTTTC TAAAAGTTTT TCCATATCAA CAGTGAAGTT AGGATTTTCA | 3300 |
| ATCACGATAC CACGATTTC AGCATGACCG ATATTTTCAA TAATTTTCAGC GTTATGAAGG | 3360 |
| TAGGTCCTGG TTGGAATACA TCCACGGTTT AAGCAGGTC CACCAAGTTC AGATTTCTCA | 3420 |
| ACAAGGGCAA CCTTACCGCC GAATTGGGCA GCTTTAATGG CTGCAACATA ACCAGCAGGA | 3480 |
| CCTCCACCAA TCACAACGAT ATCAAAAGCA TCATCGCTCT TACCATCATC GTTTGAGGTA | 3540 |
| CTTGCTACAG GTACAGGGCT AGCTTCTGGC GATGCTGCTC CAGCTGTGG GATGTTTTCC | 3600 |
| CTTTCTTCAC CAAGGTAACC GATAACTTCC GTTACAGGGA CAGTTTCACC ATCTCCTTTG | 3660 |
| AGAATGGCAA TCAAGTACCC ATCTTCTTCG GCTTCCAATT CCATGCTGAC TTTATCAGTC | 3720 |
| ATGATTTCCA AAAGGATTTC TCCTTCTTTT ACAAATTCCTC CGACTTTTTT ATTCCATTGG | 3780 |
| ACGATTTGTC CTTCTGTCTAT ATCCACGCCG GCTTTGGGCA TAATTACTTC TAAGGCCATG | 3840 |
| TCTTCCTTCC TTTATCTATA TCTTAAAAAT GAATACTCTT GCTCTTAAAT TAACATTGAG | 3900 |
| ATTGGCGTTT CAATCAACTC TTTCAAGTCC TTCATAAACT TAGCACCAGC CATACCATCT | 3960 |
| ACGACACGGT GGTCAATGGT TAATCCTAAA CTCATGATTG GGCGAATCAC AATTCACCA | 4020 |
| TTGACGACAA CTGGCTTCTC GATTGTCGAA CTGACACCAA GGATAGCTGA GTTGGGTGG | 4080 |
| TTAATAATCG GACCAAAGGA CTGAACACCA AACATTCCCA AATTACTGAT TGTGAATGTT | 4140 |
| GAATTTTGTA ACTCACTTGG AGCCAATTTA CCATCCAAGG TACGGCCAAT AACATCCTTA | 4200 |
| AAGGCTACAA CCAGTTCTGA AAGACTCATC TTCTCAGCAT TGTAACAAC AGGTGTCATC | 4260 |
| AATCCATTAT CCATCCCAAC TGCCATGGCA AGATTGACAT AGTTGTGAGT GATAATAGTC | 4320 |
| TTGCCATCTT CTGTCAATGA AGCGTTGATG TATGGGTGTT TCATAAGAGT CTTAACAACT | 4380 |
| GCAAGCGAAA GAAGGTCTGT TACAGTAGTC TTCTTCCAG TTGCTTCCAT GATTGGCTCA | 4440 |
| AGAACCTTCT TACGAAGAGC CAACATTTCA GTCATATCAA CTTCATAGTT GAGGGTGAAG | 4500 |
| GTTGGCGCAG TCAAGTAAGA TTCAACCATG CGTTGGGCAA TAACCTTACG CATTGGTGTC | 4560 |
| ATTGGAATAC GCTCGATTTT ACCATATGGT GTTACGTTAT CAGGGACTTC TTCCACTTTT | 4620 |
| TCAATCTGAG CAGGAGATTT GATGCTATCG TTTTCGATAT TTTCAAGGAAG CAGGGCCAAA | 4680 |
| ACATCCTTCT TCATGATTTT ACCACGATGA CCGGTTCTTT GGATTTCCTG CCAAGCAATG | 4740 |
| TTATGTTTCA GGGCAATTCG TTTTGCAAGT GGGCAATGC GAACCACGTT TGTGCTTTTA | 4800 |
| TAAGTTTCCA CGTCTTCTTT GTGGACACGA CCGTTTGCAC CTGAGCCAGA AACGTCGTAG | 4860 |

988

| | |
|--|------|
| AGGTTTATCC CTAAATCATC CGCTAACTTT CTAGCTGCAG GAGTCGCTCT TAGCTTGTC | 4920 |
| TCAGCCATGA CCTCTCCAAT TCTATTTATG ATACAAAGGG CGTCAAAGC GACTGAAAA | 4980 |
| TAGGAAATCG ACGATGGCTT CGATGAAGCC AAGGAGATTT ATCTTTTTTC CGATCTTTTA | 5040 |
| GCCCCGTGCTC TAATCTAAGA TATTAATGAC GAAGAGCTCT GCACCTAAAA GATACAAAGT | 5100 |
| TTCTCGTCAG CTTTATTTTA TTTACATAAC TTATCTTATG TAACCCTATT CTTTGTATA | 5160 |
| AGTTTTTCGG ATTGCATCTT TGATACTTTC AACTGTTGGA ATCATTGCAT TTTCTAGGTT | 5220 |
| TTGTGCATAA GGCATCGGCA CATCTTCTCC TGCACAACGG CGAATTGGTG CATCTAGATA | 5280 |
| GTCAAATGCT TCTGATTCTG AAATAATAGC TGAATTTCA CCGATATAGC CACTTGTTTT | 5340 |
| GTGGGCATCG TTGACCAGAA CAACCTTACC AGTCTTCTTC ACTGAGTTTA TGATGATATC | 5400 |
| CTTATCAAGC GGAACAAGGG TACGTGGGTC AACAAATTTCA ACTGAAATTC CTTCTTCTGC | 5460 |
| TAATTCTTCA GCAGCTTGAA CCACACGGCG AAGCATTTTT CCATAAGTAA CAACTGTTAC | 5520 |
| ATCCGTTTCT TGGCGTTTGA TTTACCAAC CCCAAGTGGA ATTGTGTAGT CTGGATCAAC | 5580 |
| TGGCACTTCC CCTTTTGGT TAAATTCTGA CTTGTACTCA AGTATAATAA CTGGGTGTGT | 5640 |
| ATCACGGATA GAAGACTTAA GCAGGCCTTT CATGTCCGCA GGTGTCCAG GTGCCACAAC | 5700 |
| CTTAAGTCCT GGAATGTGAG TAAACCAAGA CTCTAGAGAT TGTGAGTGCT GGGCGGCAGA | 5760 |
| GCCAACTCCG TTACCAGCTG CACAACGAAC AGTCATTGGA ACCTGACCTT TACCACCAA | 5820 |
| CATGTAACGT GTTTTAGCAG CTTGGTTGAC GATATTGTCC ATGGCAATAA CAGAGAAGTC | 5880 |
| CATGAAGGTC ATATCGACGA TTGGACGAAG TCCTGTCATG GCTGCTCCTG CTGCTGCTCC | 5940 |
| AGAGATGGCA GCTTCAGAAA TCGGACAGTC ACGGACACGT TCTGGACCAA ATTCTTCAAG | 6000 |
| CATTCCAACA GAAGTACCGA AGTCTCTCC GAAGACACCG ACGTCTTCTC CCATCAAGAA | 6060 |
| CACATTTTCA TCGCGACGCA TTTCTCAGA CATAGCAAGG ATAATGGTGT CACGGAAGGA | 6120 |
| CATTGTTTTT GTTTCCATT TATCTCTTTC TCCTTAGTCT GCGTAAATAT CTTCAAAGGC | 6180 |
| TGATTCAAGC GGTGGGAATG GGCTTTCCTC TGCAAATTTA ACAGAAGCTT CTACTGCTTC | 6240 |
| CTTTACTTGC GCTTGGATT CTTCCAATTC TTCGGCACTT GCAATGTTAT TTTCAATAAG | 6300 |
| GTAATTGCGG AGGTTTTCGA TTGGATCTTT TTGTTTCCAC AATTCCACTT CTTACACGCT | 6360 |
| ACGATATTTA CCAGGGTCAG ATGATGAGTG ACCGAGCCAG CGATAAGTTA CACTTTCAAT | 6420 |
| CAAGACTGGA CCATTGCCAC TGCGAACATG GTCCACAGCT TTCTGAAATC CTTCATAGAC | 6480 |
| ATCGATGACA TTGTTACCGT CTTGATGAA CATTCCAGGA ATTCATAAG CGGCGCTACG | 6540 |
| TTGATGGATA TGTTCATAT TGGTCATTTT CTTGATATCC GCAGAGATAC CGTAACCGTT | 6600 |
| GTTAATGCAA TAGAAAATGA CTGGCAGGTT CCAGATAGAA GCCATGTTCA CTGCTTCGTG | 6660 |

989

| | |
|--|------|
| GAAAAACACCT TCATTGGTCG CACCATCTCC AAAGAAGCAG ACAACGATTT TACCGGTATT | 6720 |
| TTGCATTTGC TGACTGAGGG CTGCACCGAC AGCGATCCCC ATACCACCAC CTACGATACC | 6780 |
| ATTGGCACCA AGGTTCCCAG CATCAAGGTC AGCGATATGC ATAGATCCAC CTTTCCCTTT | 6840 |
| ACAGGTTCCA GTGTATTAC CAAGGATTTC AGCCATCATT CCGTTGAGGT CAATCCCTTT | 6900 |
| AGCAATAGCT TGCCCGTGC CACGGTGGTT TGAGGTAATC AGATCATCTG GATTGAGAGC | 6960 |
| TAACATAGCC CCCACGTTAG CTGCCTCTTC ACCAACAGAA AAGTGCCTCA TTCCTGGCAC | 7020 |
| TTTCCCTTTC TTTACTAATT GTGCAATTTT TAAGTCCATG CGACGGATTT CTTCATCTT | 7080 |
| ACGGAACATT TCTAGCAAAA GATTTTATC TAAAGTTGAC ATCTTCTTGC CTTTCTAACT | 7140 |
| TTCTTCTTAC CTACTATTT TACCGCTTTT GGCAAATACT GTCAAAGTTT TTCTAAAAGA | 7200 |
| AATTTACAA AATAAAAAAG AAAACCCCGT GAAACAAGG GATTTTCTTG TCAAGAATAT | 7260 |
| TTTTTCACAA ACTTTTTAGC ATTTGGATTT TGCTAAAGAT TCAAATCTCT TCATAATCAC | 7320 |
| AGTTAAACGC CAACGGTAGA GCGCCCGCT CACAATCAAA CTAATAATCA AGCCGATCCA | 7380 |
| GTAAGAATAA GCTCCAAAAT CTGTTAGGGA ATCAAATAGC GTAnCACAGG GATTGCTACG | 7440 |
| CCCCAATAAC CAAGCAAACC AAGGTAAAA GGAATAACTG TATCCTTATA CCCCCGAAA | 7500 |
| ATTCCTGAA GCGGCGCCGC AAAGGTATCT GCTAACTGGA AGAAAAGACT ATAAGTTAAA | 7560 |
| AAACGCACTG TCAAATCGAT AAATTTTGGG TCGTTACCAT AAAGACTGGC CACATTTCCC | 7620 |
| CTAAAAATGT AAAGGAAGGT TAAGGTGAAG GCCGCAAAAA TGAGGGCAGT CCATCTTCCT | 7680 |
| AGACCAATAT AGGTTTTCGC ATCATCAAAT CGTTGGCTC CCACTTCATA GGAACGACA | 7740 |
| ATAGCCATAG CCGATGAGAT ACTCATAGGA AAGGCGTACA TAAGACTTGA AAAGTTCATA | 7800 |
| GCTGACTGGT GACTAGCTAT AATCAAGGGC GAAAACTTAG CCATAATCAA GCCAACCCT | 7860 |
| GAAAAGATAG CCACTTCCGC GAAGACAGTT CCCCCAATAG GCAGACCTAA ACGAACTCCT | 7920 |
| TCCTTAATTT TATCCATATT AAGTGAATT CGTTTCTCAA GGTGTAAGGC TTTGAGCTTC | 7980 |
| TCCTGTTTAA ATAAAACCAG AACAGAAATC CCAAGCAAGA CCCAGTAGGC CAAGGATGTT | 8040 |
| CCTAAACCAG CACCAGCCCC TCCAGTTCT GGAACACCAA AGGCACCGTA AATCAAGAGA | 8100 |
| TAGTTAAATC CGCTATTGAG AGGGAGTAAC AAAAGCATGA GGTACATGGA CAGTTTGGTC | 8160 |
| AAGCCCAGCG AATCCAGCAA GGAACGAATG ACGCTAAAGA GCAACAAGGG GATAATCCCG | 8220 |
| ATAGATAAAA ACCAAAGATA GCGAACCCT ACTGCCGCTA CTGCTGCTTC TAACCAATA | 8280 |
| TGATTCAAGA TTATTGGTGC CAAGAAAAGT ACCATCCCCA GCAAGACCAC AGATAGGCCC | 8340 |
| AAGGCCAAAT AAATAAATTG GTAAAAATCA GACGCAACTT CTTCTTTTTC GCCTCGACCA | 8400 |

990

| | |
|--|-------|
| AGATGGTGAC CAATGATAGG CACCAAGGCT GACACAATCC CTGTTAGAAA TGTAAGAAA | 8460 |
| GGATTCCAGA TACTGGTTGC CATAGATACA CCAGCCAAGT CCATAGTGTT GTATTGACCT | 8520 |
| GTCATTGCAG TATCAACAAA AGAGGCAGAA TAATTGGCAA ATTGGTAGAT CAGGATTGGG | 8580 |
| AAGAAATTTT TAAAAATAA TACTAACTTC TCTCGTAAAC ACTTTGTCTT ATACATACTT | 8640 |
| CTCTTTCTAT TCTGATTAT CTAACCCAAA GAGTTTCAGA CCATAGTTTT TCAAACCTAG | 8700 |
| CGGAGGTTTA TTAGATTTTG AAGTAGTATG CCAACACGCA CATGTACGAC AATAATAGCT | 8760 |
| TCTAACTAAA CCTCCGTTAT CATATTGAAC CGCATGGTCA GCTTTTTCTT TAGTTTCATA | 8820 |
| TTGAATTTTG GAACGATTAG CTGCGGGACA GTAAATTCCA CTATTAGATT TCGCTTGTCT | 8880 |
| CTCCCTACGT TTTCGAAAAT AATTCATATT CTAACCTCCTA TCAAGCTTGA TAGACGATTT | 8940 |
| GTCCCTTACA GATGGTATAT TTAACCTGCC CTTTAAAGGT TTCACCGATG AATGGTGAAT | 9000 |
| TAGCTGCTTT GGAAGCAAAA TGGGAGTCCA CAAAGCGGTC AGCCTTGGCA TCAAAAATAG | 9060 |
| TGATATCTGC TGGACCATTG TCAGCCAAGT AACCTGCTTC AAAGTTGTAA AGCTTGGCTG | 9120 |
| GGTTGTATGT CATTTTTTCA AGTAATTCCA TCAAGCTCAA CTCACCAGCT TCTACTAAAT | 9180 |
| AGGTCAAGCT GAGAGACAGG GATGTTTCTA AGCCAGTCAT ACCAGATGGC GCTTTGGTAA | 9240 |
| TATCCTCAAC ATTTTTTTCA TCTACATGAT GAGGCGCGTG GTCAGTCGCA ATAAGTGTGA | 9300 |
| TGACACCTGA TTTGAGACCT TCGATAACGG CACGACGGTC TGATTCCAAA CGAAGCGGTG | 9360 |
| GATTCACTT AGCATTGCTA CCTTGTGTTA AAAGAAGTGC TTCTGTCTTA GAGAAATGCT | 9420 |
| GTGGCGCTAC TTCTGCTGTG ACTTCTGCAC CTAACCCCTG AGCAAACCTCC ACTACTTTAA | 9480 |
| CACTTTCTTC CTTAGACAAA TGCTGGATGT GAACATGGGC TTTAGTTGCA TAGGCAATCA | 9540 |
| TGACATCAG CGCCATCATA GCGTACTCAG CCACCCAGT AGCACCGCAG ATATGGAAAT | 9600 |
| GTTCTCTAGC AATATTTTCA TTAAAGCCAA GAACACCGTT CAAACCTGGA TCTTCCTCAT | 9660 |
| GAAGGCTGAT AAAGGTATTG AGTTTTTTGG CTTCCTCCAT GGCTTCCTTG ACAATCTTAC | 9720 |
| TGCTCTCAAG CGGAATACCG TCATCAGAGA AACCAACCGC ACCAGCTTCT AAGAGTGCCT | 9780 |
| TAAAGTCAGT CAAGTTTTTA CCATTAAAGT TTTTAGTAAT GGTGCAACT GTCTTGACAT | 9840 |
| TAATCTTCTC TTTGGCAGCT GACTGGAGAA CTGCTTGCAA AGTCTCCACG TCTGAAATGG | 9900 |
| TTGGACTGGT ATTAGCCATC ATGACGACAG TAGTAAACC ACCTGCAGCG GCTGCTAGGG | 9960 |
| CACCAATATG AATGTCTTCT TTATGTGTTT GACCAGGTTT ACGGAAATGA ACATGAATAT | 10020 |
| CGACCAAGCC AGGAGCAACC ACAAGACCAG TAGCATCAAT CGTTTCTGCT CCTTCTTCCG | 10080 |
| TGATCTCAGA CGCAATTTTG ATAATTTTCC CATCTTGAAC TAAGACATCA CAAACTTGAT | 10140 |
| CCAAACCAGA CTTGGGATCC ATTACACGAC CATTTTGTAT TAGTAGCATC TGCTTTCTCC | 10200 |

991

| | |
|--|-------|
| TTTATTCATA GAAATCAACT TGGGTATCCA ACAATTTATC CCCATCATAA ACAAACCTGG | 10260 |
| CTGAAAAGAA GGGTTTATCC TCTAAAAGCC ACTCAACAAA GGTGTGGTCA CCTTCCCAAG | 10320 |
| TCGGCTTGCT CAAAACCTCA TCATAGGGAA CCCATTCTAG CGTCCCCTCA TTGCAGTCAA | 10380 |
| TCAAGTCGCC CTCAAACCTCC GTCACCTTAA AAACATAGGT GTACCAGTCT AAATCTGGTG | 10440 |
| TAAATTCAGG AAAAGTGATG ACACCTTTTA GAACTGGCTT GGCTTTGAGC CCTGTTTCTT | 10500 |
| CAAGGATTTC ACGCGCCGCG CATTCTGGG GCGTCTCTCC TCTCTCTAGC TTACCACCCA | 10560 |
| CACCAATCCA TTTCCCTTCA TGGACATCAT TGGGTTTCTT ATTACGATGG AGCATGAGCA | 10620 |
| GTTCCTTCCC ATTATCAATG TAGCAAATCG TCGCTAACTG AGGCATATTT TCTCCTTATC | 10680 |
| TAAGCCAATC GATTGGCTCT TGTCTGTCT CTTTAAAGAA TGCATTGGCC TTGGAAAAGG | 10740 |
| GCTTGGAACC CAAAATCCT CTATAAACCG ACAAAGGACT TGGATGGGCT GATTGATAA | 10800 |
| TCAAGTGATG AGGATGGTA ACTAATGCCT TCTTCTTACG TGCATAAGCT CCCCAGAGTA | 10860 |
| CAAAAACGAC TGGTCTATCT AGATGATTGA CCACCTGAAT CACAGCATCA GTAAAAGGCT | 10920 |
| CCCAGATTG ACCAGCATGA CCATTGGCCT GTCCAGCAGG AACAGTCAA CAAGCATTAA | 10980 |
| GAAGCAAGAC TCCTTGCTCA GCCCAAGCTG TCAAATCATG AGATTCTTA ACTCCGATAT | 11040 |
| CATCTGACAA TTCTTTCAAG ATATTTTGCA AGGATGGTGG AGCTGGGATA GAGTCAGGTA | 11100 |
| CAGAAAACT CAAGCCCTGC GCTTGACCTG GTCCGTGATA GGGGTCTTGC CCTAGAATTA | 11160 |
| CCACCTTAAC TTCTTCAAGC AGTGTGTGCA AGAGAGCCTG AAAAACCTTT TCCTTGGGTG | 11220 |
| GATAAATAAT CCCCTGAGAA TAGACCTGCT CCATAAACTG ATTGATTTTC CCGAAATAAC | 11280 |
| CCTCAGGTAA TTGCGCCTTA ATCAAAGCAT GCCAAGACGA GTGTTCCATA GCCGACTCGG | 11340 |

(2) INFORMATION FOR SEQ ID NO: 148:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 12127 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 148:

| | |
|---|-----|
| AAAAAATAGA CTTGTTAGAC TATAAATGTA GTAAGCCTAC ACAAGAAAAA TACATAGAGA | 60 |
| TAAAGGTGAT TATTATGAAA TTCAAAAAAA TGCTTACTCT TGCAGCCATT GGCTTATCAG | 120 |
| GATTTGGGCT TGTTGCCTGT GGCAATCAGT CAGCTGCTTC CAAACAGTCA GCTTCAGGAA | 180 |
| CGATTGAGGT GATTTCACGA GAAAATGGCT CTGGGACACG GGGTGCCTTC ACAGAAATCA | 240 |

992

| | |
|--|------|
| CAGGGATTCT CAAAAAAGAC GGTGATAAAA AAATTGACAA CACTGCCAAA ACAGCTGTGA | 300 |
| TTCAAAATAG TACAGAAGGT GTTCTCTCAG CAGTTCAAGG GAATGCTAAT GCTATCGGCT | 360 |
| ACATCTCCTT GGGATCTTTA ACGAAATCTG TCAAGGCTTT AGAGATTGAT GGTGTCAAGG | 420 |
| CTAGTCGAGA CACAGTTTTA GATGGTGAAT ACCCTCTTCA ACGTCCCTTC AACATTGTTT | 480 |
| GGTCTTCTAA TCTTTCCAAG CTAGGTCAAG ATTTTATCAG CTTTATCCAC TCCAAACAAG | 540 |
| GTCAACAAGT GGTACAGAT AATAAATTTA TTGAAGCTAA AACCGAAACC ACGGAATATA | 600 |
| CAAGCCAACA CTTATCAGGC AAGTTGTCTG TTGTAGGTTT CACTTCAGTA TCTTCTTTAA | 660 |
| TGGAATAAT AGCAGAAGCT TATAAAAAAG AAAATCCAGA AGTTACGATT GATATTACCT | 720 |
| CTAATGGGTC TTCAGCAGGT ATTACCGCTG TTAAGGAGAA AACCGCTGAT ATTGGTATGG | 780 |
| TTTCTAGGGA ATTAACCTCT GAAGAAGGTA AGAGTCTCAC CCATGATGCT ATTGCTTTAG | 840 |
| ACGGTATTGC TGTGTGGTC AATAATGACA ATAAGGCAAG CCAAGTCAGT ATGGCTGAAC | 900 |
| TTGCAGACGT TTTAGTGGC AAATTAACCA CCTGGGACAA GATTAAATAA AATGTTTGCT | 960 |
| CCATAAATCT CTAAAGAGAT GCAGACGTTT CATCGTACAA TAAGATAAAG AAGGCAAGTA | 1020 |
| GGGAGGTGTC GTATCTCCCT TACTTTCTTC ACTAGAAAGG ACAAGATGTG ACAAACAAG | 1080 |
| CCTTCAAAGA AGCAGTTTTT AGGGCAATTT TTTTCATGAG TGCAACAGTA GCTGTTGTAG | 1140 |
| CTATTTTGCT AATCTGTTTC TTTATTTTTA GTAATGGCTT ACCTTTCATA GCTAACTACG | 1200 |
| GCTTTGCCCG TTTTATTTA GGCAGTGATT GGTGCCAAC GAACATTCCG GCAAGCTATG | 1260 |
| GTATTTTACC AATGATCGTT GGTTCCTTAT TAATTACCTT AGGAGCGATT GTGATTGGGG | 1320 |
| TGCCAACAGG CATCTTGACA TCGGTGTTTA TGGTTTATTA TTGTCCAAAG CCGTCTATG | 1380 |
| GCTTCTTAAA ATCAGCTATC AACTTGATGG CAGCCATTCC ATCTATTGTT TATGGTTTTT | 1440 |
| TCGGCCTACA ATTATTGGTG CCTTGATTA GAAGCTTTTT AGGAAATGGC ATGAGTGTC | 1500 |
| TAACCGCTTC GTTACTATTA GGAATAATGA TTTTGCCAAC CATTATCAGT TTGTCAGAAT | 1560 |
| CTGCTATCCG AACAGTTCCC AAAACGTATT ATTCTGGTAG CTGGCTCTA GGAGCTAGTC | 1620 |
| ATGAACGGAG TATTTTATAGT GTCATCTGCG CAGCTGCGAG ATCTGGTATT TTATCAGCAG | 1680 |
| TTATTTTAGG AATCGGTCGC GCAGTAGGTG AAACCATGGC AGTTATTTTG GTGGCAGGCA | 1740 |
| ACCAGCCGAT TATCCAAGT GGACTCTTTT CAGGAACCAG AACCTTAACA ACCAATATTG | 1800 |
| TTCTGGAAT GGCTTACGCA TCAGGTCAGC ATAGGGAAGC CCTTATTGCA ACCTCAGCAG | 1860 |
| TTCTCTTTTT CCTTATCTC TTGATTAATG CCTACTTTCG CTAATTGAAA GGAAATCAT | 1920 |
| CTTATGAGTA AATACCTGCT AAAACTTCTC GTTTATTGTT TTTTCTTTT AACCTTTGGC | 1980 |
| TCTCTTTTT TAATCATGGG TTTTATCCTC ATCAAAGGCT TACCTCATCT AAGTCTATCC | 2040 |

993

| | | | | | | | |
|------------|-------------|------------|------------|------------|-------------|------------|------|
| CTCTTTTCTT | GGA | CTTATAC | TTCTGAGAAC | ATTTCCCTTA | TGCCAGCGAT | TATTTCCACC | 2100 |
| GTATTTCTGG | TCTTTGGTGC | TCTTCTTTTA | GCCTTGCCCA | TAGGGATTTT | TGCTGGTTTT | | 2160 |
| TATCTTGTGG | AATATACAAA | AAAAGATTCC | CTTTGTGTTA | AAATCATGCG | ATTGGCCTCA | | 2220 |
| GATACCTTAT | CTGGGATTCC | TTCCATFGTT | TTTGGTCTGT | TTGGCATGCT | CTTCTTTGTA | | 2280 |
| GTCTTCTTAG | GTTTTCAATA | CTCTCTGTTA | TCAGGAATCT | TAACCTCAGT | TATCATGGTG | | 2340 |
| TTGCCAGTCA | TTATTCGCTC | AACAGAAGAA | GCCCTTTTAT | CTGTTAGTGA | TAGCATGCGT | | 2400 |
| CAAGCAAGTT | ATGGACTTGG | GGCAGGTAAG | TTACGGACTG | TTTTTAGAAT | TGTTCTACCA | | 2460 |
| GTGGCCATGC | CAGGTATTTT | AGCTGGAGTG | ATACTAGCTA | TTGGCCGTAT | CGTTGGTGAA | | 2520 |
| ACAGCTGCCC | TCATGTATAC | ATTAGGTACC | TCTACCAATA | CGCCAAGTAG | TCTCATGTCT | | 2580 |
| TCAGGCCGTT | CTCTAGCCCT | ACATATGTAT | ATGCTGTCAA | GTGAGGGGCT | ACATGTCAAT | | 2640 |
| GAAGCCTATG | CTACCGGCGT | GATTTTGATT | ATTACTGTTT | TAATGATAAA | TACTCTATCA | | 2700 |
| AGCTTATTAT | CTCGAAAAC | TGTGAAAGGA | GCTTCCTAGT | ATGGGAACAT | TTTCAGTCAG | | 2760 |
| ACACCTAGAC | TTATTTTACG | GGGATTTTCA | AGCCTTAAAA | AATATTTCGA | TTCAATTACC | | 2820 |
| AGAAAGACAG | ATTACTGCCT | TGATAGGCCC | ATCTGGTTGT | GGCAAATCAA | CTTTTCTAAA | | 2880 |
| AACCCCTAAC | CGGATGAACG | ATTTGGTTCC | TTCTTGCCAT | ATTGAAGGCC | AAGTCCTCTT | | 2940 |
| AGATGAGCAA | GATATTTATA | GTAGCAAATT | CAACCTTAAT | CAGCTACGTA | AGCGTGTAGG | | 3000 |
| GATGGTTTTT | CAACAGCCTA | ATCCCTTTGC | CATGTCTATC | TATGATAACG | TGGCTTATGG | | 3060 |
| CCCAAGGACA | CATGGTATTC | GAGACAAAA | ACAATTAGAT | GCCTTAGTGG | AGAAATCTTT | | 3120 |
| AAAAGGGGCA | GCCATTTGGG | AAGAAGTCAA | AGATGATCTT | AAAAAGAGTG | CCATGTCTCT | | 3180 |
| ATCTGGCGGT | CAGCAGCAAC | GCCTTTGCAT | TGCGCGAGCT | TTAGCAGTAG | AACCTGATAT | | 3240 |
| TCTGTTAATG | GATGAGCCGA | CTTCAGCCTT | AGACCCTATC | TCCACTTTAA | AAATTGAAGA | | 3300 |
| CCTCATTCAG | CAACTAAAA | AGGATTATAC | GATTATCATT | GTTACCCATA | ACATGCAACA | | 3360 |
| AGCTTCACGT | ATTTTCAGATA | AAACTGCTTT | TTTCTTAACA | GGAGAAATTT | GCGAATTTGG | | 3420 |
| AGATACCGTT | GACGTGTTTA | CCAATCCAAA | AGATCAGCGC | ACAGAAGACT | ATATTTTCAGG | | 3480 |
| ACGGTTCGGA | TAAGGAAGGA | AAAACCTATG | AGAAATCAAT | TTGACTTAGA | ATTGCATGAA | | 3540 |
| TTAGAACAAT | CCTTTTATAGG | ACTAGGGCAA | CTTGTCTTGT | AAACAGCTTC | AAAAGCCTTA | | 3600 |
| CTGGCCTTAG | CCTCCAAAGA | CAAGGAGATG | GCAGAGCTAA | TTATCAATAA | GGATCATGCT | | 3660 |
| ATCAACCAAG | GTCAAAGCGC | TATCGAATTG | ACCTGTGCCC | GTTTGTGGC | CTTGCAGCAG | | 3720 |
| CCACAAGTGT | CTGACCTTCG | ATTTGTGATT | AGCATCATGT | CTTCTTGTTT | AGACCTTGAA | | 3780 |

| | | |
|--|------|--|
| 994 | | |
| CGTATGGGAG ACCATATGGC AGGCATTGCC AAAGCTGTTT TGCAACTAAA AGAAAATCAA | 3840 | |
| CTAGCCCCCTG ACGAAGAACA GTTACACCAA ATGGGTAAAT TATCCCTCAG CATGCTAGCC | 3900 | |
| GATTTATTGG TTGCCTTTCC TTTGCACCAA GCCTCAAAAG CTATTAGTAT TGCTCAAAAA | 3960 | |
| GATGAACAGA TTGACCAATA TTATTATGCC TTATCAAAGG AAATCATTGG ACTTATGAAA | 4020 | |
| GACCAAGAAA CCTCAATTCC CAATGGAAC CAATACCTTT ATATCATAGG GCATCTGGAA | 4080 | |
| CGCTCGCTGA TTACATTGCT AACATTTGTG AACGCCTAGT CTACCTAGAA ACAGGAGAAC | 4140 | |
| TAGTGGATTT GAATTAATTC AACTAATCCT TAAAAGAGAA GAGTACGATT AAGTACTCTT | 4200 | |
| TTTTATGGTT GTAAAAAGT TCATTTGACC AATTAAAGCA GTGTAGATAG TGAGGAGTTG | 4260 | |
| TTTCAATTCT ATCGTGAACG AGGGAATGCT GAAAACCTTA TCAAAGAAAG GAAAGCAGGA | 4320 | |
| TTCTTTGGGG ATAAGACAGA TAGTTCGACC ATGATTAAGA ATGAAGTACG TATGATGATG | 4380 | |
| GGCTGTCTGG CTTATAATCT CTACCTCTTT TTAAAGCAGC TAGCTGGTGA TGAAGTAAAG | 4440 | |
| TCCTTGACTA TCAAGCGTTT TCGACGTCTC TTCCTTCATA TTGCCGAAA ATATGTCTCT | 4500 | |
| ACTGCTAGAC GACATATTCT CAAATTCTCA AGTCTATACG CCTATTCAAA ACAGTTTCAA | 4560 | |
| GCCTTATTTG ATACAATCTG CCAGATAAAT CTGATACTCC CTGTTCCATA TAGAGCTAGA | 4620 | |
| GGGCAGGGGA AAACATGCCT AACAGAATAA GTCACCTTAT TTTAAAAATC GAGCATCAAA | 4680 | |
| CCAAGGGAGG AGTCTGCCCT TTTTtagGAA AAAATCAAGA CAAATCTCCT CAATTATGTC | 4740 | |
| TCGAACATCA GAAATTAAGC AAAATCACCA GAAGGACAGT ATTTCAACTA GCTTTTCTGG | 4800 | |
| TAATTTTGA ACTGTGTAGT TCGTTAGTGC CAGATATGAA TAATTGGGA TGATAAATCT | 4860 | |
| TTCTTCCTCA GGTAGCCTAT CATAATACTC TTCAAAAATC TTATCAAAAA CACTCTCTTT | 4920 | |
| CTTTTGGGCG ATAGTTTCAT CTTCGTATGT AGGAGTCCTC ATCAAGAAAT ACTTCAATTC | 4980 | |
| TAGGTATTCC TTATCCAACT CTATATAACT TGGCATCAAC TTGTAATCTT CAACCCCAA | 5040 | |
| ACGTTCAAGCA ATATATTTTA ACTTTGTTAG TATTGGTCTG GATTCTCCAT TTTCAATTCT | 5100 | |
| AATTAATTGA CGGATACTTA ATTCAGACTC ATCACCACAA AATTCTGAAC GACTGATTI | 5160 | |
| TTTAGCCAAA CGTAATCTTT TAATTTTTTC GCCAAACTCT CGCAACCTAC AAGAACTTCC | 5220 | |
| TGAGTTGTTT ACCTCTATTA TAAGCATATA CTGAATCAAA CTATCTATCA GATTTCTTCT | 5280 | |
| CACTTTAACT AAAGACTAAG AGTTTATCCC TTCGTCTCGG TTTTGTGTA TTTTCCACC | 5340 | |
| ATACCCAGT AATGCAAGTG CAAAATCCCC TAGAATATGA TAGAATAAGA GAAAGAACTC | 5400 | |
| TATCAAGGAG GAAATCATGG AAAAACAAC CGTCGCCGTC TTGGGGCCTG GTTCTTGGGG | 5460 | |
| AACCGCCCTT TCACAAGTCT TAAATGACAA TGGACACGAG GTACGTATTT GGGGAAATCT | 5520 | |
| TCCCGAGCAA ATCAATGAAA TTAATACACA CCATACTAAT AAGCACTACT TTAAAGATGT | 5580 | |

995

| | |
|--|------|
| CGTTCTAGAC GAAAATATCA TTGCCTACAC CGACTTAGCA GAAACATTGA AAGATGTGGA | 5640 |
| TGCGATTTTG TTTGTTGTCC CAACAAAAGT GACACGACTT GTTGCCCAAGC AAGTTGCACA | 5700 |
| AACCTTGGAC CATAAGGTTA TCATCATGCA CGCATCAAAG GGATTAGAAC CTGATAGCCA | 5760 |
| TAAACGATTA TCAACCATTG TTGAAGAAGA AATTCTTGAA CATCTCCGTA GTGATATCGT | 5820 |
| CGTTGTTTCA GGGCCTAGTC ATGCAGAAGA GACCATTGTG CGTGACCTAA CTTTAATAAC | 5880 |
| TGCTGCTTCT AAAGATTTAC AAACAGCTCA ATACGTTTCTG AAGCTATTTA GTAATCACTA | 5940 |
| CTTCCGACTT TATACCAATA CGGATGTTAT CGGGGTTGAA ACTGCTGGTG CTCTTAAAAA | 6000 |
| TATATTGCT GTCGGTGCTG GAGCTTTACA TGGTCTTGA TTTGGTGATA ATGCTAAGGC | 6060 |
| AGCCATCATC GTCGAGGTT TAGCAGAAAT CACCCGCTTA GGGGTAGCAC TCGGGGCCAG | 6120 |
| TCCATTGACC TATAGCGGCT TATCTGGTGT GGGAGATTTG ATCGTAACGG GAACTTCCAT | 6180 |
| CCACTCTCGT AACTGGAGAG CTGGAGATGC TCTCGGACGA GGAGAATCCC TAGCTGATAT | 6240 |
| AGAAGCTAAT ATGGGCATGG TAATCGAAGG AATTTCACG ACTCGAGCAG CCTATGAACT | 6300 |
| AGCCCAAGAA CTTGGAGTCT ATATGCCCAT TACACAGGCT ATTTACCAAG TTATTATCA | 6360 |
| CGGAACCAAT ATCAAAGATG CCATTTATGA CATCATGAAC AATGAATTTA AAGCAGAAAA | 6420 |
| TGAGTGGTCT TAACCCTCTA TAGAAAGGAT TTTTATGACA TCAAAAAGTTA GAAAGGCAGT | 6480 |
| CATCCCTGCT GCTGGACTAG GAACTCGATT TTTACCAGCA ACCAAGGCC TTGCCAAGA | 6540 |
| AATGTTGCCA ATCGTAGACA AACCAACTAT CCAGTTTATC GTGGAAGAAG CTCTCAAATC | 6600 |
| AGGTATTGAA GATATTCTAG TTGTCACTGG TAAATCAAAA CGTTCTATTG AGGACCACTT | 6660 |
| TGATTCAAAC TTCGAATTGG AATATAACCT CAAAGAAAAA GGGAAAACAG ATCTTTTGAA | 6720 |
| GCTAGTTGAT AAAACAACG ACATGCGTCT GCATTTTATC CGCCAAACTC ATCCACGCGG | 6780 |
| TCTCGGAGAT GCTGTTTTGC AAGCCAAGGC TTTCTGCGGA AATGAACCTT TTGTCGTTAT | 6840 |
| GCTTGGTGAT GACTTGATGG ATATCACAGA CGAAAAGGCT GTTCCACTTA CCAAACAAC | 6900 |
| CATGGATGAC TACGAGCGTA CCCACGCGTC TACTATCGCT GTCATGCCAG TCCCTCATGA | 6960 |
| CGAAGTATCT GCTTACGGGG TTATTGCTCC GCAAGGCGAA GGAAAAGATG GTCTTTACAG | 7020 |
| TGTTGAAACC TTTGTTGAAA AACCAGCTCC AGAGGACGCT CCTAGCGACC TTGCTATTAT | 7080 |
| CGGACGCTAC CTCCTCACGC CTGAAATTTT TGAGATTCTC GAAAAGCAAG CTCCAGGTGC | 7140 |
| AGGAAATGAA ATTGAGCTGA CAGATGCAAT CGACACCCTC AATAAAACAC AACGTGTATT | 7200 |
| TGCTCGTGAG TTCAAAGGGG CTCGTTACGA TGTCGGAGAC AAGTTTGGCT TCATGAAAAC | 7260 |
| ATCCATCGAC TACGCCCTCA AACACCCACA AGTCAAAGAT GATTGAAGA ATTACCTCAT | 7320 |

996

| | |
|--|------|
| CCAACTTGGA AAAGAATTGA CTGAGAAGGA ATAACAAAAT CATTTATATA AAGATTAGCC | 7380 |
| ACACATAAAT TAAGTAAATT CTCTACTTGA ATCTACCTAT TTAATAAAAA CTAATGAAAA | 7440 |
| CGCTATACTT GTATTTGTTT TTTCATTAAA ATAAGAGTAG AATAAATTAG TATAGTAAAA | 7500 |
| CAAAAAAGCA CCGAATCGGT GCGCACTTTT TCAAGTTGTG TACGGACAAA GCCTTATTTT | 7560 |
| AACTTTGCTA TGTGTGTTCT AATGGTTCCA AAATAATAAA TAATTTTAAA TTTGACTTAA | 7620 |
| CTGTTGGAGT AGTCATGGTT AAATTAAATC AACCGAGCCG AACATAAGTT GTTTAATTTT | 7680 |
| GTGGAAGCTA TTAATAAAAA TATAATAAGG GAGAAAAGATA GGTGTAATTT TAATTTTAAA | 7740 |
| GTAATTGCGG ACACTATCAA AGAAAAAGAT TATGGAGAAC AAATTTGTAG AATTTATCGA | 7800 |
| AAACAATAAA AAAGTAATCA TTTCATCAGT TGCAGTTGGT GTTGTATTGG TATTAGGGTT | 7860 |
| TGGATGGTAT TCATATAACC AACACAAGC AGAACAACAA GCAAAAATTG TACAATTAGA | 7920 |
| AAAAGATAGC AAATCAGACA AAGAACAAGT TGATAAACTA TTTGAATCAT TTGATGCATC | 7980 |
| TTCAGATGAA TCTATTTCTA AATTAAGA ACTATCTGAA ACTTCACTTA AAACCGATGC | 8040 |
| AGGTAAAGAC TATCTTAATA ACAAAGTCAA AGAATCATCT AAAGCAATTG TAGATTTTCA | 8100 |
| TTTGCAAAAA GGTTTGGCTT ATGATGTTAA AGATTCAGAT GACAAATTTA AAGATAAAGC | 8160 |
| AACTCTTGAA ACAAATGTAA AAGAAATTAC AAAACAAATT GATTTTATCA AAAAAGTTGA | 8220 |
| TGAAACTTTT AAACAAGAGA ATTTGGAAGA AACTCTTAAA TCTCTAAATG ATCTTGTTGA | 8280 |
| TAAATATCAA AAACAAATCG AACTTTTGAA GAAAGAAGAA GAAAAAGCTG CTGAAAAAGC | 8340 |
| TGCTGAAAAA GCAAAGGAAT CTTCTAGTCA AAGTAATCT TCTGGTAGTG CTTCTAATGA | 8400 |
| GTCTTATAAT GGATCTTCCA ATTCAAATGT AGATTATAGT TCATCTGAAC AAATAATGG | 8460 |
| ATATTCAAAT AATTATGGCG GTCAAGATTA TTCTGGTTCA GGAGATAGTT CAACAAATGG | 8520 |
| TGGATCATCA GAACAATATT CATCTAGCAA TTCAAACAGC GGAGCAAATA ATGTCTACAG | 8580 |
| ATATAAAGGC ACTGGTGCTG ACGGCTATCA AAGATACTAC TACAAAGATC ATAATAATGG | 8640 |
| AGATGTGTAT GATGACGATG GAAATTACCT TGGGAACTTT GGTGGCGGCA TTGCAGAACC | 8700 |
| TAGTCAACGC TAATAACTAT TTTAGAGCTG TGTGTTTCG AATGGTTCCA AAACACATTA | 8760 |
| AAAGCTACTC ATTTTTTAAG TAGCTTTTTT CTTATTCAAG TTTACATATT ATACTCAATG | 8820 |
| AAAATCAAAT TCAAACCACG TCAGCATCGC CTTACCGTAG GTATGGTTAC TGACTTCGTC | 8880 |
| AGTTTCATCT ACAACCTCAA AACCATGTTT TGAGCTGACT TCGTCAGTTC TATCTACAAC | 8940 |
| CTCAAAGCAG TGCTTTGAGC AACCTGCGGC TAGCTTCCTA GTTTGCTCTT TGATTTTCAT | 9000 |
| TGAGTATTAG TCGTCACAAT CCCATTCCCT TGTAGAAAAG CAAATGGCG AGTCCTACGA | 9060 |
| ACAAGACTAC CGCTCCTAAT CTCTGGCTGG TGTATACAT CCGTTTTCTT CCTCTAACTG | 9120 |

997

| | |
|---|-------|
| GAAAGATAAC TGCTAGAAAT GCGCCACCAA CTGCACCACC GATATGGCCT GCTAGGCTGA | 9180 |
| TTCCTGGAAT CAGAACACTT CCAATAATGT TAACCACAAA AAGTGTCAGA TAGGATTGCC | 9240 |
| CTAGCTGTTG GATATAAGGA TTGCGAGTTG CATAGCGAAG AACATAATC GCGGCAAATA | 9300 |
| GCCCATAAAG AGAGGTAGAG GCGCCTGCTG CTAAGGATTT AGGACTAAAT AAAAAACAA | 9360 |
| AGAGATTGCC CATCATTCCT GATAAAGAT AGAGAAAGAA AACTGCTTA GAACCGAAAA | 9420 |
| TCTCCTCTAC CTGCCTTCCA AGATAATAAA GTGAAAGCAT ATTAACAATG AAATGTTCCC | 9480 |
| ACCCAATATG AACAAAAATG GCAGACAAGA GACGCCAAC CTGCTCGGGA AAGAGGCGAA | 9540 |
| TAGCTGGCCC ATACATGGCT CCAATCGAA ATAATGTATC TGCCCTGTCA AAGTTTCCGC | 9600 |
| CTGCAGTGAC CAACATTAGT AAAAATACCA AGGCCGTCAC TAAGAGGAAG AACTCGTCA | 9660 |
| CAGGGTAACG TCTATCAAAG ATTTCTTCA TCAATTAATA CCTCTGAAC AGGAATATCA | 9720 |
| TGGTTTTAG GTATAAAGTC CTGAATTTGA CAAGGATATA TCGTACTCAA AGTACGACCA | 9780 |
| GAAAAATGTT CCAGATAGCG GTCATAATAG CCTCCACCGT ATCCTATCCG ATATCCTTTC | 9840 |
| GTCGTAAGG CCAGACCAGG AACATGAATC AAATCAATCT GAGATGCATC CACCCTTCC | 9900 |
| AAATCTCCCT GTAGCTCCAG TAAGGCAAAG AAAGTTTTTA CCAACTGTTG CGGATCATAG | 9960 |
| ACCACAAAGT CCATGCGCCC CTTGGGATAA GTTTTGGGTA TTAACCTT CTTGCCGTCC | 10020 |
| TTCAGCGCCT GCTCAATCAG TTCCTGCGTT TGAAACTCAT GAGAAAAGA GAGGTAGGTT | 10080 |
| GCGATGACCT TGGCTTCTTG ATAAAAGGGG TGTTGTAAAA GCCGCTCGGT TAAAGCTTGG | 10140 |
| TCTATAGCCT GTTTTGTCTT TTGAGATATA GCCTTCATTT CATGCAAGAC TTGCTGCGT | 10200 |
| AATTCCGATT TCATAGACAA GCCCTCTATT CTGCTGCCTT CTTTTTCAGG AACTAGACA | 10260 |
| CCGCAGCCAC CCCAATAGCT AAGACTTCTT CCTTAGGACT CATTTGAGGG TGATGAAGAG | 10320 |
| CGTAGGGACT ATCGATACCT AGCCAAAACA TCACGCCATC AACCTTTGAA AGGAGATAAC | 10380 |
| CAAAGTCCTC GCCTGTCATA GCAGGTTGGA TATCAATCAA CTCGATTCCG TCTTTTTCGT | 10440 |
| CAAAGAAGTC CATCAGTTCA CGCGCCAAGG CTGGATTGTT CTCAACAGGT AGGTATCCAC | 10500 |
| CTTGTTTGAG TTCCACTTCG ACTTCCATAT CAAAGGCAGC TGCAACCCCT TCTGCAACTG | 10560 |
| TTTTTACCCT CTTTTCACC AAGAGACTCA TGTCCTGTGT CAAGGCACGA ATAGTTCCAT | 10620 |
| GTAAAAAAGC TGTGTCTGTG ATGACATGTG TGGTGGTTCC AGCTTGAAAA ACGCCGAAGG | 10680 |
| TCACCACTGC TCCCTCGATT GGGTTGACAT TGCGGCTAAC AACTGACTGC ACTTGGGTCA | 10740 |
| CAAAGTAACT AGCCGCCACC AAGGCGTCAT TGGCTTCATG AGGAAAAGCT GCGTGGCCAC | 10800 |
| CTTTCCTTTT GAAACGGATC TTCACCTCGC AAGTTCCTGC AAAGAGTGTA TGAGTATTAG | 10860 |

998

| | |
|---|-------|
| TCGCAATCTG GCCGACTTTC AAATCTGGAC GAACATGGAG ACCATAGAAT TGATCTGGCA | 10920 |
| ACCAATCTCC AAAAGCACCG TCCTCATACA TGAGCATACC ACCAGCTTCA TTTCTTCAG | 10980 |
| CAGGCTGAAA TAGAAAGAGC AGATTATTCT TGGGTTGCTC CTCAAGGGCG CGCTCAAGAC | 11040 |
| AGCCTAAGGC AATGGTCATA TGAAAATCAT GGACACAGGC ATGCATGCGA CCTTGGTGTT | 11100 |
| GAGAAGCAAA AGGTAGACCT GTTTGTTCGA CGATAGGCAG GCCATCAATA TCTGTCCGCC | 11160 |
| AACCAATGGT TCGCTCCGGC TGAATCCCT GCAGGTAGAC CAAAATCCCT GTCCGCCAAG | 11220 |
| TACGAATTTG AACAAAATCC TTGCCGTAG TCAATTTCTC AATCACATCC AGCAAATAAG | 11280 |
| CCTGAGTCTT GAACTCCTCC AAGCCAATCT CTGGAATCTG GTGTAAATCT CGTCTAGTCT | 11340 |
| GAATCAAATC TAACATCTAT CTGTCCTCCG ATATAGCAGA AAGAGGCTGG AAAAAGGGTT | 11400 |
| CCGCTCTTTT TTTACTTTTA CAATTACAAG GTACGAAGCG CATCCTCTAG CGCTGTTTTT | 11460 |
| TGTTGAGTTT GGGCATCAAT TTCTTTGATA ATACGAGCTG GAACACCTGC TACTACCAG | 11520 |
| TTTTCTGGGA CATCTTGGGT AACAATAGCT CCGCTGCGA CAACTGAACC ACTACCGATT | 11580 |
| TGGACTCCTT CGATAACCAC TGCATTAGCA CCGATAAGAA CATTGTCTCC GACACGGACT | 11640 |
| GGTTCAGCAC TAGCTGGCTC AATCACACCT GCCAAAACCTG CACCTGCACC AACGTGGCTA | 11700 |
| TTTTTTCCAA CGATGGCAGC GCCACCAAGG ATGGCACCCA TGTCAATCAT GGTTCAGCA | 11760 |
| CCGATTTTCT CACCGATATT GATAACAGAT CCCATCATGA TAACAGCATT GTCACCAATT | 11820 |
| TCCACCTGGT CACGGATAAT CGCACCTGGC TCGATACGAG CGTTGATAGC ACGCTTATCT | 11880 |
| AGCAAAGGAA CTGCAGAATT ACGAGCATCT TGCTCGACAA CATAATCTTG ATTTTCTACC | 11940 |
| AAACCTTCAA GAAGCGGAGC CACATCCTTC CAGTCTCCGA ATAGGACATT TCCTAGTTTG | 12000 |
| ACAACAGAGC TAGGCACAGC AGTTGCGAGT TGCCCCCTCAA AGGTTACTTT GAACTGGTT | 12060 |
| TTCTTTTCTCAG CATTGGCGAT AAATGGGATA ATTTCTTGAG CGTTCATTTT TGTAGCAGTC | 12120 |
| ATAGGTG | 12127 |

(2) INFORMATION FOR SEQ ID NO: 149:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 12566 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 149:

| | |
|---|-----|
| CCATCCTTCT GTTGATGTGA CAGGAATGAT GATAAATCAA CCAGTAGCTA GTCGGAAGA | 60 |
| GGTGACAGAG GCTTTGAGTC ACTTGGCGGT AGAGCACAAT AGTCTCATTG CTCGTCGAAT | 120 |

999

| | |
|--|------|
| CGTTGAGCCA AATGAAGCTG GAGAAACACG CTTTACCTAT GCCACTTATG GTGAGGGAAA | 180 |
| GCTTCCAGAA GGTCTGACCA TTTCCTCCAA GGAGAGTGCA GAAACGAGTG ATTTATTAGG | 240 |
| GTCTTACTTG ATTGTATCAG GAAGTTTGA TGGAGTGAGC TTACAGACCA CCTTGAAAGA | 300 |
| GCTTGGTTAT CAAGGCTTTG TTTCGAATGG AGAAGATCCA TTTTCGATAG TCTTACTATT | 360 |
| GACGGCCACC CCTATGGTGC TACTGAGTTT AGCTATTTT CTGCTGACCT TTATGAGTCT | 420 |
| GACCCGTGATT TATCGGATCA AATCCCTTCG TCAGGCAGGG ATTCGCTTAA TAGCTGGTGA | 480 |
| GAGCTTGT TT GGAGTTGCTC TCAGACCAGT GTTAGAAGAT GTGAGACAGC TTATCTGCTC | 540 |
| AGTGCTGGTA TCCAGTCTTT TGGGATTGGG GATTCTCTGG TATCAAGGTG CCTTGT TTAT | 600 |
| GGCAACGGTG CAACTGGTCA TCATTGCTCT TCTACTTTAT GGATTGACCT TGGCAGGGAT | 660 |
| TTCTACCTTA CTAAGTGTG TCTATCTACT TGGTTTACAG GAAATAGTC TGGTGGATCT | 720 |
| ATTGAAAGGG AAACCTCCCTC TCAAACGTAT GATGACATTG ATGATGGTGG GGCAACTCTT | 780 |
| AGCTGTATTG GTGGTCGGAT CGAGTGCGAC AGCTCTCCTA CCCCCTACC GTGAAATGCA | 840 |
| GGAAATGGAG AGAGCTAGCA ATAAATGGAG CCAGTCCTCA GACCGTTACC GTCTATCCTT | 900 |
| TGGTTGGTCT AGTGCAATTG CCGATGAAGA AGGAACGCGT AAGGATAATC GTGAGTGGCA | 960 |
| GACATTTACT GAAGAACGGT TAGCCAATAC AGACTCTTTT TATATTATGA GCAATGTGTA | 1020 |
| CAATTTCTCA GATGGAGCAG AAGTGGACCT AGATGGCAAT CGTCTCAGTG ACTACACACC | 1080 |
| GTCAGGGAAT GTTATCTATG TCTCACCGCG CTATCTGATA GAAGAAAAGA TTACCGTTTC | 1140 |
| TTCAGAGTTT ATGGACAAGA TGCAAACTT GTCTGAGGGA GAGTTTGGGC TGATCTTGCC | 1200 |
| TGAGAGCTTG CGAGAGCAGT CTGTCTACTA CCAAGGATTG TTTACAGATT ACCTGCAAAA | 1260 |
| CTTTTCATCT GAAAGTGTAG AAGTGACGAG TCAGAAACAC TACCTCCCAC AGGTAAGGCT | 1320 |
| AGCTTTTACA GAAACAGGAC AGGAACGTTT CCTCTATAAT GATGGGTACA AGACAACACG | 1380 |
| CCAGTACCTA AAAGATCCGA TTATTGTAGT TCTAACGCCG CAAGCGACTG GAACAAGACC | 1440 |
| TGTTGCAGGG ATGTTGTGGG GAACTACGGC TAATAGTGCC TTGAAACTAG ATCGATATGG | 1500 |
| AGACAGCATC ACAGCTCTAA AAGAGAAAGG TCTGTATCAC AAGGTTTCTT ACTTGGTAAA | 1560 |
| AAGCCAGCTA TTTTGTGCCA AGGTACTAAA TGACAAACGG GTGGAGTTT ACTCTCTCCT | 1620 |
| TATTGGGACG ATTTTGACCC TGTCTACGGC TATCTTGTTA TTTGATTCCA TGAATCTTCT | 1680 |
| CTATTTTGAG CAGTTCAGAC GGGAACTTAT GATTAAACGT CTTGCTGGTA TGACAATCTA | 1740 |
| TGAGCTTCAT GGCAAGTATT TACTGGCGCA AGGAGGAGTT CTCTTGCTTG GCCTAGTCCT | 1800 |
| ATCTAGTATT TTGACAAGAG ATGGTTTGAT TAGCGCTCTA GTGTAGCTT TGTTTACGCT | 1860 |

| 1000 | | | | |
|-------------|------------|------------|------------|-----------------------------|
| TAACGCCCTC | TTGATTTTAG | TAAGGCAGGA | CAAAAAAGAA | GAAGCTGGTA GCATGGCAGT 1920 |
| ATTGAAAGGA | AAATAAGATG | ATTGATATTC | AAGGATTGGA | AAAGAAATTT AATGACCGCG 1980 |
| CGATTTTCCTC | TGGTTTGAAT | CTCAAGCTGG | AGAAGGGCAA | GGTTTATGCC TTAATCGGAA 2040 |
| AGAGTGGAAG | CGGAAAGACG | ACGCTGCTGA | ATATCTTGGG | AAAGCTAGAA AAGATAGATG 2100 |
| GTGGAAGGGT | TCTCTATCAG | GGGAAAGATT | TAAAAACCAT | TCCCACCTCGT GAGTATTTTC 2160 |
| GAGACCAGAT | GGGCTATCTC | TTTCAAAATT | TCGGCCTCTT | AGAAAACCAA TCAATCAAAG 2220 |
| AAAATTTGGA | TTTGGGTTTT | GTTGGTCAGA | AAATCTCAA | AGTAGAACGT TTGAAAGGC 2280 |
| AAGTGGGGGC | TTTAGAAAAA | GTTAATCTAG | GGTATTTGGA | TTTAGAACAA AAAATCTATA 2340 |
| CTTTATCTGG | GGGAGAGGCC | CAACGAGTTG | CCCTTGCTAA | GACTATTTTG AAAATCCAC 2400 |
| CCTTGATTTT | GGCAGATGAA | CCAACAGCAG | CTCTTGATCC | TGAAAATTCA GAGGAGGTTA 2460 |
| TGAATCTCTT | GGTGGATTTG | AAAGATGAAA | ATCGAATTAT | CATCATTGCG ACCCATAATC 2520 |
| CCCTAGTCTG | GAATAAGGCT | GATGAAATCA | TTGATATGAG | GAAACTTGCT CATGTGTGAA 2580 |
| AAAATCCGTA | TTGCGAGGGT | ATCTGATTAT | CCTAGTGCCA | GAGGTGGTTT AGAAGATATC 2640 |
| CTCATCATGG | AAAATATGAC | CAATCATCTC | CTTTTGGTTC | AAATCCGAGT GCATGGCTAT 2700 |
| TTGCTTGATT | TTGCTAGTAT | TGAAGGGCAA | AGGCAAAAGC | ATTATCGTTT GAAAAATTTA 2760 |
| CCTCAGACGG | TTGAACTGAC | AGTGGATGAT | GTGGAGGAGG | ATGTGGATTT GACCCACCT 2820 |
| GAAAATCGAA | GTTATCAAGA | AGCTGATTTT | TTTGAACGCA | TGTTTCGAGA GAACTGCTAA 2880 |
| GGCCACTTTT | AAAGATTTCC | AAGACTATCT | TTCTTCATGA | GGAAAGATAG TTTTTTGGTA 2940 |
| TGATTTTCAT | TCCCAAATA | CAAGGGGAAT | GTGTTACAAT | AGTAGTAACA GATAATAGAA 3000 |
| AAGAGAATAG | ATGAGAATTG | CAGATTATAG | CGTGACCAAG | GCAGTGCTGG AGCGTCACGG 3060 |
| TTTTACCTTT | AAAAAGTCCT | TTGGGCAAAA | TTTTTTGACG | GATACCAATA TCCTTCAAAA 3120 |
| AATTGTGGAT | ACGGCTGAAA | TTGATGATCA | GGTCAATGTC | ATCGAAATCG GGCCAGGTAT 3180 |
| TGGTGCCTTG | ACAGAATTTT | TGGCTGAGCG | TGCAGCCCAA | GTCATGGCTT TTGAGATTGA 3240 |
| CCACCGTTTG | GTGCCAATTT | TGGCAGATAC | CCTGCGTGAT | TTTGATAATG TGACCGTAGT 3300 |
| TAACGAAGAT | ATTCTCAAGG | TTGATTTGGC | GCAACATATC | CAGAATTTTA AAAATCCTGA 3360 |
| CCTGCCAATC | AAGGTAGTGG | CTAATTTGCC | TTACTACATC | ACGACGCCTA TTCTCATGCA 3420 |
| CTTGATTGAG | AGTGGCATTC | CTTTTGTGA | GTTTGTGGTC | ATGATGCAGA AAGAAGTAGC 3480 |
| GGACCGCATT | TCAGCCCAGC | CTAACACCAA | GGCTTACGGT | AGCTTGTCTA TCGCCGTGCA 3540 |
| GTATTACATG | ACAGCCAAGG | TTGCCTTTAT | CGTGCCTCGT | ACGGTCTTTG TGCCAGCGCC 3600 |
| AAATGTGGAT | TCAGCCATCT | TGAAAATGGT | GCGTCGTCCA | GAGCCAGCCG TAGCAGTAGA 3660 |

1001

| | |
|--|------|
| AGATGAGAAC TTTTCTTTA AGGTTTCCAA GGCTAGTTTT ACCCATCGCC GCAAGACCTT | 3720 |
| GTGGAATAAC TTGACAGGTT ACTTTGGTAA GACTGAAGAG GTCAAGGACA AGCTGACCAA | 3780 |
| GGCTTTGGAC CAGGCAGGCT TGTACCAAG TGTGCGTGGG GAAGCTCTCA GCTTGGCAGA | 3840 |
| ATTTGCCGGT CTAGCAGACG CACTTAAAGG GCAAGGACTC TAAGATGCAG GGACAAATCA | 3900 |
| TTAAAGCCTT GGCAGGTTC TACTATGTGG AGAGTGATGG CCAGGTTTAT CAAACACGCG | 3960 |
| CGCGTGGGAA TTTCCGTAAA AAAGGCCATA CCCCTTATGT TGGGGACTGG GTAGATTCTT | 4020 |
| CTGCCGAGGA AAATTCAGAA GGCTATATCC TCAAAATTCA CGAACGGAAA AACAGTCTGG | 4080 |
| TTCGTCCGCC TATTGTCAAT ATCGATCAAG CTGTAGTAAT CATGTCCGTC AAGGAACCTG | 4140 |
| ATTTTAACAG CAATTTGCTG GATCGTTTCT TGGTTCTTTT GGAGCACAAAG GGCATCCATC | 4200 |
| CCATTGTCTA TATTTCCAAA ATGGATTTGT TGAAGATAG GGGAGAACTG GATTTTACC | 4260 |
| AGCAGACCTA TGGTGACATC GGCTATGACT TTGTGACCAG TAAAGAGGAA CTCCTGTCTT | 4320 |
| TGTTAACAGG CAAGGTTACG GTCTTTATGG GGCAGACAGG TGTGGGAAG TCAACTCTTC | 4380 |
| TCAATAAAAT CGCACCAGAC CTCAATCTTG AAACGGGAGA AATTTAGAC AGTCTAGGTC | 4440 |
| GCGGTCGCCA TACCACTCGA GCTGTTAGTT TTTACAATCT CAACGGGGGT AAAATCGCAG | 4500 |
| ATACACCAGG ATTTTCATCC TTGGACTATG AAGTATCAAG GGCTGAAGAC CTCAATCAGG | 4560 |
| CTTTCCCA GATTGCTACT GTTAGCCGAG ATTGTAAGTT CCGTACTTGT ACCCATACCC | 4620 |
| ATGAGCCGTC TTGTGCCGTC AAACCAGCTG TTGAAGAGGG TGTATTGCA ACCTTCCGTT | 4680 |
| TTGACAATTA CCTGCAATTC CTTAGTGAAA TTGAAAATCG TAGAGAAACC TATAAAAAAG | 4740 |
| TCAGCAAAAA AATTCCAAAA TAAGGAGAAA CCTATGTCTC AATACAAGAT TGCTCCGTCA | 4800 |
| ATTCTGGCAG CAGATTATGC CAACTTTGAA CGTGAAATCA AACGTCTAGA AGCAACTGGG | 4860 |
| GCAGAATATG CCCATATCGA TATCATGGAC AGTCATTTTG TACCGCAAAT CAGTTTGGT | 4920 |
| GCAGGTGTGG TCGAGAGCCT TCGTCCTCAT AGTAAGATGG TTTTCGATTG CCACTTGATG | 4980 |
| GTGTCAAACC CTGAGCATCA TCTGGAAGAT TTTGCGCGTG CAGGTGCAGA CATCATCACT | 5040 |
| ATCCATGTAG AAGCAACGCC TCATATTCAT GGCGCCCTCC AAAAAATTCG TTTACTCGGA | 5100 |
| GTTAAGCCTT CAGTCGTTAT CAATCCTGGC ACATCAGTTG AAGCCATCAA GCACGTCCTT | 5160 |
| CATCTAGTTG ACCAAGTTTT AGTCATGACG GTTAATCCAG GTTTTGGTGG GCAAGCCTTT | 5220 |
| CTGCCAGAAA CCATGGATAA GGTCCGTGAG TTGGTTGCTC TTCGTGAGGA AAAAGGTTTG | 5280 |
| AACTTTGAAA TCGAAGTGGA TGGTGGGATT GATGACCAAA CTATTGCTCA AGCCAAAGAA | 5340 |
| GCCGGTGCGA CTGTTTTTGT AGCAGGTTCC TATGCTTTTA AGGGAGAAGT CAATGAGCGA | 5400 |

1002

| | |
|--|------|
| GTACAAACTC TCAGAAAACA ACTGGACTAG GGTTCAGTTC TTTGCAGGCG GAAACCGCGG | 5460 |
| TCATTATCGG ACAGATTTTG ATGCTTTTGT TGGGGTGGAT CGAGGCTCGC TCTGGGTCTT | 5520 |
| GGAAGAAGAC TTACCTCTTG CTCTAGCAGT CGGAGATTTT GATTCTGTGA CGGAAGAAGA | 5580 |
| GCGACAGGTG ATTCAAAAAG GTGCCCAGTA TTTTGTCCAA GCACGACCAG AAAAGGATGA | 5640 |
| TACAGATCTG GAATTGGCTC TCTTAACCAT CTTTGAACAA AATCCTCAGG CTCAGGTCAC | 5700 |
| TATTTTCGGT GCCTTGGGTG GCCGTATTGA CCATATGTTG GCCAATGTCT TTCTGCCTAG | 5760 |
| CAATCCTAAG TTGGCACCTT ATATGCATCA AATAGAAATT GAGGATGGGC AAAACTTGAT | 5820 |
| TACTTATTGT CCAGAAGGAA TCAGTCAGCT AGAACCTCGT TCAGACTACG ACTATCTAGC | 5880 |
| CTTTATGCCA GTTCGGGATA GCCAGCTGAC TATTCTTGGA GCCAAGTATG AGTTGACAGA | 5940 |
| GGAAAATTTT TTCTTTAAAA AAGTGTACGC TTCTAACGAA TATATAGATA GGGAAAGTGC | 6000 |
| GGTAACTTGC CCAGATGGTT ATGTGGTCGT ACTGCATAGC AAGGACAGGA GGTAGGATGG | 6060 |
| AAAGTTTACT TATTCTATTA TTAATTGCCA ATCTAGCTGG TCTCTTTCTG ATTTGGCAAA | 6120 |
| GGCAGGATAG GCAGGAGAAA CACTTAAGTA AGAGCTTGGA GGATCAGGCA GATCATTTGT | 6180 |
| CAGACCAGTT GGATTACCGC TTTGACCAAG CCAGACAAGC CAGCCAGTTA GACCAAAAAG | 6240 |
| ATTTGGAAGT GGTGTGCAGC GACCGTTTGC AAGAAGTGC GATTGAATTG CACCAAGGTC | 6300 |
| TGACCCAAGT CCGTCAAGAA ATGACAGATA ATCTCCTCCA AACTAGAGAC AAGACAGACC | 6360 |
| AACGCTCTCA AGCCTTGCGA GAATCAAATG AGCAACGTTT GGAACAAATG CGCCAGACGG | 6420 |
| TCGAGGAAAA ACTAGAAAAG ACCTTGCGA CACGCTTACA GGCTTCCTTT GAGACAGTTT | 6480 |
| CTAAACAACCT GGAGTCTGTC AATCGTGGCC TTGGAGAAAT GCAGACAGTT GCCCGTGATG | 6540 |
| TCGAGAGCTCT TAACAAGGTT CTCTCTGGAA CCAAGACGCG AGGGATTCTG GGAGAATTGC | 6600 |
| AACTGGGGCA AATTATTGAA GACATCATGA CACCTGCCCC GTACGAACGA GAATACGCAA | 6660 |
| CGGTTGAAAA CTCTAGTGAA CGAGTGGAGT ATGCCATCAA GTTACCCGGA CAAGGCGACC | 6720 |
| AAGAATACGT CTATCTGCCA ATTGACTCTA AGTTTCCACT GGCAGATTAT TACCGCTTGG | 6780 |
| AAGAAGCCTA TGAGACAGGT GACAAGGATG AGATTGAACG CTGTCGTAAG TCACTCCTAG | 6840 |
| CAAGCGTCAA GCGCTTTGCT AGGGATATTA GGAACAAGTA CATAGCACCA CCTCGGACGA | 6900 |
| CCAATTTTGG AGTTTGTGTT GTTCCGACAG AAGGTCTCTA CTCAGAAATC GTCCGCAATC | 6960 |
| CGGTCTTCTT TGATGATTTG AGACGGGAAG AACAGATTAT TGTTGCAGGA CCAAGTACCC | 7020 |
| TATCAGCCCT TCTTAACTCC CTATCAGTTG GTTTCAGAC CCTTAATATC CAAAAGAGTG | 7080 |
| CCGACCATAT CAGCAAGACT CTTGCCAGTG TCAAGACCGA GTTTGGCAAG TTTGGTGGTA | 7140 |
| TTCTGGTCAA GGCACAAAAA CATCTCCAAC ATGCCTCTGG CAATATTGAT GAATTATTAA | 7200 |

1003

| | |
|--|------|
| ACCGTCGTAC CATAGCTATC GAGCGGACGC TCCGTCACAT TGAGTTGTCA GAAGGTGAGC | 7260 |
| CTGCGCTTGA TCTACTCCAT TTTCAAGAAA ATGAGGAAGA ATATGAAGAT TAGTCACATG | 7320 |
| AAAAAAGATG AGTTATTTGA AGGCTTTTAC CTAATCAAAT CAGCTGACCT GAGGCAAACCT | 7380 |
| CGAGCTGGGA AAAACTACCT AGCCTTTACC TTCCAAGATG ATAGTGGCGA GATTGATGGG | 7440 |
| AAGCTCTGGG ATGCCCAACC TCATAACATT GAGGCCTTTA CCGCAGGTAA GGTGTGCCAC | 7500 |
| ATGAAAGGAC GCCGAGAAGT TTATAACAAT ACCCCTCAAG TCAATCAAAT TACTCTCCGC | 7560 |
| CTGCCTCAAG CTGGTGAACC CAATGACCCA GCTGATTTC AAGTCAAGTC ACCAGTTGAT | 7620 |
| GTCAAGGAAA TTCGTGACTA CATGTCGCAA ATGATTTTCA AAATGAAAA TCCTGTCTGG | 7680 |
| CAACGGATTG TCCGAAATCT CTACACCAAG TATGATAAGG AATTCTACTC CTATCCAGCT | 7740 |
| GCCAAGACCA ACCACCATGC CTTTGAAACG GGCTTGGCCT ATCATACGGC GACCATGGTG | 7800 |
| CGTTTGGCAG ACGCTATTAG CGAAGTTTAT CCTCAGCTCA ATAAGAGCCT GCTCTATCGC | 7860 |
| GGGATTATGT TGCATGACTT AGCTAAGGTC ATCGAGTTGA CGGGGCCAGA CCAGACAGAG | 7920 |
| TACACAGTGC GAGGTAATCT TCTTGGACAT ATCGCTCTCA TTGATAGCGA AATTACCAAG | 7980 |
| ACAGTTATGG AACTCGGCAT CGATGATACC AAGGAAGAAG TCGTTTGTCT TCGTCATGTC | 8040 |
| ATCCTCAGTC ACCACGGCTT GCTTGAGTAT GGAAGCCCAG TCCGTCCACG CATTATGGAA | 8100 |
| GCAGAGATTA TCCATATGAT TGACAATCTG GATGCAAGCA TGATGATGAT GTCAACAGCT | 8160 |
| CTTGCTTTGG TGGATAAAGG AGAGATGACC AATAAAATCT TCGCTATGGA TAATCGTTCC | 8220 |
| TTCTATAAAC CAGATTTAGA TTAATAATTT AAGAAAAATG AGCATTTTTT AGGATAAGAA | 8280 |
| TGTTTCGTTTT TTTATGTGAA TATGGTATAA TAAGTAAAG AAAAAATGA ATACTCTTCG | 8340 |
| AAAATCTCTT CAAACTAGGG TAGTATCGCC TTGTCGTATG TATATATGCA GGTATATTAC | 8400 |
| AGGGTTTGTG AGTTCTATTG ACAATCTCAA AACAGTGTTT TGAACCACCA GCGACCAGCT | 8460 |
| TTCTAGTTTG CTTTGTGATT TTTTGAATAA AAATGGAATA GGAAATAGAA ATGAAATTAA | 8520 |
| GAAGAAGTGA TCGGATGGTT GTCATTTCCA ACTATTTGAT TAATAATCCT TATAAACTAA | 8580 |
| CTAGTCTCAA TACTTTTGCT GAAAAGTATG AGTCTGCTAA ATCATCCATC TCAGAAGATA | 8640 |
| TCGTCAATTAT CAAACGCGCC TTTGAGGAAA TTGAAATCGG TCATATCCAG ACAGTGACTG | 8700 |
| GGGCTGGCGG AGGTGTCATC TTCACACCGT CTATTTGAG TCAGGATGCT AAGGAAATGG | 8760 |
| TTGAAGACTT GCGTACCAAG TTGTCAGAAA GTGACCGTAT CTTGCCAGGT GGTATATCT | 8820 |
| ATCTGTCTGA TTTGCTTAGC ACACCAGCCA TCTTGAAAAA TATTGGTCGT ATTATTGCCA | 8880 |
| AAAGCTTTAT GGACCAAAAA ATTGACGCGG TTATGACCGT AGCAACTAAG GGTGTGCCAC | 8940 |

1004

| | |
|--|-------|
| TTGCAAATGC AGTTGCCAAT GTCCTCAATG TCTCTTTTGT CATTGTGCGC CGTGACCTGA | 9000 |
| AAATTACCGA AGGTTCAACT GTTAGCGTCA ACTATGTTTC AGGTTCAAGT GGTGACCGTA | 9060 |
| TCGAGAAAAAT GTTCCTTTCA AAACGTAGTC TTAAGGCAGG CAGCCGTGTC TTGATTGTGG | 9120 |
| ATGACTTCTT GAAAGGTGGC GGAACGGTCA ATGGTATGAT TAGTCTCTTG CGCGAGTTCTG | 9180 |
| ACTCAGAACT GGCAGGTGTA GCGGTCTTTG CGGACAATGC CCAAGAAGAA CGTGAAAAGC | 9240 |
| AGTTTGACTA CAAGTCACTC TTGAAGGTAA CCAATATTGA TGTCAAGAAC CAAGCCATCG | 9300 |
| ATGTTGAGGT TGGCAATATC TTTGACGAAG ATAAATAAGA GATAGAACTA AAGGTTGGAA | 9360 |
| CGATTGTCCC AGCCTTTCTT TGCAAACAGA ATAGAAGGAA GCTTATGAAA ACACCATTTA | 9420 |
| TCAATAGAGA AGAGTTAGAA GCGATTGTTG CCGAGTTCCC GACTCCCTTT CACTTGTATG | 9480 |
| ATGAGAAGGG GATTCTGTAG AAGGCAAGAG CCGTCAACCA AGCTTTTTCG TGGAACAAGG | 9540 |
| GCTTTAAGGA ATATTTTGCA GTTAAGGCTA CTCCAACCTC AGCTATTTTG AAAATCTCC | 9600 |
| AAGAAGAAGG TTGTGGTGTG GACTGCTCTA GTTATGTAGA GCTTTTGATG AGCCATAAAC | 9660 |
| TGGACTTTCT GGGTTCTGAG ATTATGTCTT CTTCCAACAA CACGCCAGAC AAGGAATACG | 9720 |
| CCTATGCACG TGAATTGGGT GCGACCATTA ACTTGGATGC CTTTGAAGAT ATTGAACATC | 9780 |
| TGGAGAGAGT AGCAGGCATT CCAGAAATCA TCTCTGTGCG TTATAATCCT GGAGGCGTTT | 9840 |
| TTGAACTGGG GACAGACATT ATGGACAATC CTGGGGAGGC TAAGTTTGGC ATGACCAAGG | 9900 |
| ACCAGCTCTT TGAAGCCTTT GCTATCTTGA AGGAAAAAGG AGCCAAGACT TTTGGGATTC | 9960 |
| ACTCCTTCCT AGCGTCCAAT ACCGTGACCC ATCTCTATTA TCCAGAGTTG GTCGTCAGC | 10020 |
| TCTTTGAACT GGCTGTTGAA ATCAAGGAAA AGTTGGGCAT TTCGCTAGAC TTTATCAATC | 10080 |
| TTTCTGGCGG TATTGGTGTT AATTATCATC CAGACCAGGA GCCGAACGAT ATCGCCTTGA | 10140 |
| TTGGTGAGGG AGTTCGTAAG GTGTATGAAG AGGTTCTTAC GTCAGCAGGT CTTGGTCAGG | 10200 |
| TCAAGATTTT CACCGAATTG GGTCTTTTGA TGCTGGCACC TCACGGTGCT CTAGTCACAA | 10260 |
| GAGTCACTCA TAAGAAGGAA ACCTACCGTA CCTATCTAGG TGTGGATGCC TCAGCAGTCA | 10320 |
| ACCTCATGCG TCCAGCTATG TACGGAGCTT ACCATCATAT TAGCAACGTG ACCCATCCAG | 10380 |
| ATGGACCAGC TGAAGTGGTA GATGTGGTCG GTTCACTCTG TGAAAACAAT GATAAAATTG | 10440 |
| CAGTTAATCG CGAACTGCCT CATACAGAAA TCGGTGATTT GCTGGTCATT CATGATACAG | 10500 |
| GTGCCCACGG ATTTTCAATG GGCTACCACT ATAATGCCAA ATTACGTTCT GCGGAAATCC | 10560 |
| TCTATACCGA AGAAGGTAAA GCGGTCAAA TCCGCCGTGC AGAGCGCCCT GAGGACTATT | 10620 |
| TTGCAACCTT ATATGGCTTC GATTTTGAAG AATAATCTGA TAATAGATTG AAAATGAAAT | 10680 |
| TGAAAAACAG ATTGCTTTCT AAAAAATAGG CAAAAATCTT GTTTTCTCTT CAAGTCGTGA | 10740 |

1005

| | |
|--|-------|
| TATAATAAAA CTATAAACG TTTTCAAGGA AGGTAACGAT ATGTCTGAAG AAACAATTGA | 10800 |
| TTATGGACAA GTGACAGGAA TGGTGCATTC GACAGAAAGC TTTGGGTCAG TAGATGGGCC | 10860 |
| TGTTATTTCG TTTATTGTCT TTTTGCAGGG CTGTCACATG CGTTGCCAGT ATTGCCACAA | 10920 |
| CCCAGACACT TGGGCTATGG AGTCCAATAA GTCACGTGAA CGGACGGTAG ATGATGTCTT | 10980 |
| GACAGAGGCC TTGCGCTACC GTGGTTTCTG GGGAAATAAG GGTGGGATTA CAGTCAGTGG | 11040 |
| AGGAGAAGCT CTCTGCAGA TTGATTTCCT GATTGCTCTC TTCACCAAGG CTAAGGAACA | 11100 |
| AGGAATCCAC TGTACCTTGG ACACCTGTGC TCTTCCTTTC CGTAATAAAC CACGTTACCT | 11160 |
| TGAGAAGTTT GACAACTCA TGGCTGTCAC TGACTTGGTT CTTTGGATA TCAAGGAAAT | 11220 |
| CAACGAAGAA CAGCACAAGA TTGTCACTAG CCAAACCAAT AAAAATATCT TGGCTTGTGC | 11280 |
| CCAGTATCTA TCAGATATTG GAAAACCTGT CTGGATTCGC CACGTGCTAG TTCCAGGATT | 11340 |
| GACAGACAGA GATGATGACT TGATTGAACT TGGTAAGTTC GTCAAGACCC TCAAAAATGT | 11400 |
| TGATAAGTTT GAAATTCTAC CTTATCACAC CATGGGTGAG TTCAAGTGGC GTGAACTTGG | 11460 |
| AATTCATAT TCCCTCGAAG GAGTCAAACC ACCAACAGCA GATCGCGTCA AGAACGTAA | 11520 |
| ACAACATCATG GATACCGAAA GTTATCAAGA TTATATGAAA CGTGTACATG GATAGAAAAG | 11580 |
| AAGCCTGATG GAAACATCGG GCTTTTGACT TGCAAAAAGA CTTAGCAAAT CAGCTAAGCC | 11640 |
| TTTTTCTTCT TATCTCGAAC GTTGTTTTCC AGCGTTGCGA TTTTGTGTT TTTTCTTGCT | 11700 |
| TGTGATAGCA GTTGGTTGTT CAGGGGTAAC GTCTTTTCGT CCACTTGGTT TAGAGAAAGC | 11760 |
| ACTTGCTTTT GGTGGGTCT TGGCTAGTTC TTCACGGACT TTTTTCGAA GTTTTGGACG | 11820 |
| AACGATATAG TTGACGATAA ACTGTTGGAG AATCATCATG AAACCACCGA CAACCCAGTA | 11880 |
| AAGTGTGACA CTAGCTGGTG AGAAGAGGGA GAAGACGACG ATCATGAGTG GGCTCATGTA | 11940 |
| AATCATTTTC TTGATTGTT CTCTTGCAT TTCATCTTCT ACTCCGTGAA GTGAAAGGAG | 12000 |
| CGATTGAAGA TAGTAAAGGA CACCAGCACA GGCAACCAA ATCATACTTG GAGAACCTAG | 12060 |
| AGGAATGCCT AGGTAGCTTG CTTGAGCAAC CCCTTCAGTA TGTGGGCAG CAAAGTAGAT | 12120 |
| AGCAGAGAAG AAAGGCATTT GAAGGAGGAT AGGGAAACAT CCTACACCGC CAAACATGCT | 12180 |
| GATACCGTGC TCTTTTGTAG CAGCAAAGAG AGCTTGTGG GCTTCGAGTT TTTCTTCTTG | 12240 |
| AGTAGTCGCT TCTTTGAGAC GCGTTTGGTG TGGCTCAAG ACGTGCTTGA GGGCGTTCAT | 12300 |
| CTTTTCAGAG TGAAGCGTTG CCTTCCATGA TTGGTAGATA CCAAGTGGTA AGATAATCAA | 12360 |
| GCGTACGATA ATGGTTACGA TAATGATAGC GACACCAAAG CCTAGACCTT TATCAGTAGC | 12420 |
| GAAGTACTTG ATGGCTTCAG CCATAGGCGC TCCGATCGTA TTCCAAATAA ATCCTGTTGG | 12480 |

1006

CTGACCTGTG GTTTTATCGA CATTGACACA GCCAGTCAAG ACAAGCAACA TAGCCACTCC 12540
CATAGCCGAG AGTGCAAAAT CGGGGT 12566

(2) INFORMATION FOR SEQ ID NO: 150:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5238 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 150:

TGACACTCTG TAGGATTGTC GTTAATTGAT TGCTCGTACT CTCTACAATA ACCACCAAAG 60
TAAAAACGAC ATAGAAAGAT AGCATCAGCT GTAGCCATAG CGCCTTTGAC ACCTTCTGGA 120
TGATTATGAG TTACCTCTGC AGAAAGACTC GTAAGTCCTC TAGATGATGG CCATATACCA 180
GTTTTCGCAT AAAAACCACA GTCCATGATC CAAGCACATG GAGAAATACG CATAGCTGAT 240
CCATTCCTCA AGCTATTATA AGGCTCACGG TTATCGCTGT TTAGCCATGC ATTAAACCGA 300
GCACCGTAAT CAGCATTCGG ATACATTCTG CCATATTTCT TCATCGCGTC AATGAAGTCA 360
TCTTTTGTGTC CACCATTTCAT AATTGCTTCT GCAACAGCAC AGGTCATAAC CGTGTCATCT 420
GTAAAAAGC AGTCCTTCCG AAATAAAGGA AAGTCCTTTG TTTTGATATT GTTCCATTCTG 480
TAAACAGAAC CGACAATATC TCCAATAATT GCTCCAAGCA TCAGATTTCCT CCTTGTTTCAT 540
TTTGATGCTT TTTATATTGG TTATCTACCA TATTTATTTT AGAAAATAAC ATCCTGTGTGG 600
ATTTTAAAAA TTTTATTTT TTCAAATAG GGTTTTACCA TTTCTTTCCA CCTAGCTCTA 660
TGAAAATTGA TTGATTTTAA AGGAGATAGG CCATAATTTC CCAATGCATA ACCATCATTT 720
ACTTCAACAA CAAGTGTTCT GCCATCGCGA GTAACACCGA TATCTAGTCC ATAAGCTATT 780
GGCGCATCTT TCCAACATGA TATCGCTTCA TCAATTACAC TTGCATCAAA TTGTGCATGA 840
TAATCACCTG TATAGGGTCG AACATCTAAT ACGCGACCAT CTAACACAAA ACAACGCCAT 900
TCAGCTATGA ATTCTACAAC CTCCTAATC CATATAGGAT AGTCGAAAGG TAGACCAATA 960
CCTATTAAAT CATGGGTTC ATTAACAAC CTTCAGTAA AGACTTTTGA ACCAGCTTTA 1020
GGCTTAATAA ATTTTCCCA ATTATCAGGT ATATTCACAA TCTCTCCTAA AATACCAGCA 1080
TAAATCTTTC GACCATAAAA CTCTTTAAGC TCAATAGGAT AGTCATGAAC CGGAACGTTT 1140
AAGCCCATCA TTTTATAGTAA TGCTCTAGTC TCCATTATAT AATCTACAAC TATATCTTCA 1200
CTTGTTAACT CTTTATTTT AGAAAAAGAT TGATATAAAA TAACTTCTTC TCCTTGTAAG 1260
TAGGCACCTA CTTGAGCATT GTATTATTA ATTGAAACCT CACTTGGTAA TTTACTTTGT 1320

1007

| | |
|--|------|
| CTAATATAAA CAACCATTTTC ATCACTCCTA TATCACTAGT GTTACACCAA TTTGTAAAAA | 1380 |
| ATAATAGCAA TTTTGCTCTT ATTTTTTTGA GTAAATAGCC CCCATAATAT CATCGAAATA | 1440 |
| ATCAACGGTA TTTAGGAGTA ATTCAATAAC CTGGGACTTT GTTAGTCGCA TTCCCCTTCT | 1500 |
| ATCTCTAGCA TCTTCTACTA AATTTTCAAG TTTCTCTAGA TTTTATCAT CCAAGCTAAT | 1560 |
| CATTATTCTA TTTTATCGG TTGCCATTTT CATCACCTCA AGTTAATTCT ATCACAGGTG | 1620 |
| TAACACTAGT GTCAACTGGC TTTTATAATA CATTAGTTTA AAAGTGGAGA GGATTTTAA | 1680 |
| CACAGTAACT TTAAATCTTT GGTATTAAAA AATTTTCACA ATATTTATAG AAATAAAATC | 1740 |
| TGTCTCAAAT CAGTTATCAA ATCTAGTATA AATTATGAGC GGCTACTCTA ATACTTTCCC | 1800 |
| TCTAAACAAG AAAAAGACTT ACACTCAAGG GTTTTCTTCC CCCCCTTCGT TATAACGTTT | 1860 |
| TGACTCTTTT ACTAGCAAAG GTATATACTC ACAAGGAACT TTGGTTGACT ATGAATCTC | 1920 |
| TCCAACTTCT TCTTTAACAT ATCCTTCTAC ATCTTCAATC TCTACAAACA TTGGGTCTAA | 1980 |
| GTGACACAAG AAATGCCAAA CTTGATCCCC TTTTCTCTG TAAAGAATCG CTTACCGTC | 2040 |
| TTCACTTCCG AAAAAGCTTC TGTCGATTTT ATATCCGCGG CTTTCTAAGA AGTCTTTTGC | 2100 |
| TTTACGATAG TTCGTTTCTC TTGTTTCGAC ATAGGCTTTA ACTTCATGGT TGTTAACGAC | 2160 |
| ATATGCATCA ATTTTTGAAT ATCCTTCGAT CACTCTATCA TTTTGTAGGG ATAAATTTGA | 2220 |
| AATCTCTTTC CAAATAATGT TTACATTTTC CTCAGGATCG AACATAAATT TAGATAAAGG | 2280 |
| AACAATATTT CCGTTAAAAA TAATTTCCAT ATAATCCGGT ATGTTTTTAG GATTAAATA | 2340 |
| CTCCACTTCA AAACCATCTT CTGTTTCCAG AGTGTATCCC GGGATTTGAG CTACAAAGGC | 2400 |
| TTTCCCATCT TCTATGGAAT CAAATGCTAC TAAATCTTTA GAATAATCAT TTTGGTACAA | 2460 |
| TTCCAATATA ACCATCGATA ATCTCTCCAT TTTCATTATC AGGCTAATGT AAATAAGCAC | 2520 |
| GTCACCTGAC CAATTCAGGC TCTCTGTATC ATCTCATCAT ATTTCCCTACT TACTTTACGA | 2580 |
| GTCTTATACC CAGAACACAC CTTATCGACC TTCGGTCTCA CCTCGTCGCA TTGGCTGAAC | 2640 |
| ATCTACTTTT ACTTTGCTGA TGCTTCAACT CGTACAAGCA GTGATACCGC CTCAGCGTGA | 2700 |
| TGCGTCAGTG GGA CTCAAAA GGTTCGGGGA ACCTTTTGAG GATTAACTAC GTTTCTCTAA | 2760 |
| TAAACTTACA CATTCAACTT GTTCATCATT GTCCAAACCT ATGTTGAGAT TTTCTTCTAT | 2820 |
| AATTGGTAGC TTAAGAGTAA TGGATTTTAG CCATTGTCCG TTAGATTGTT TTTCTTCATA | 2880 |
| AACTTGAATT TCAGAAATCA AAGCTGAAAT TAACTGCCTA CGCTCTACAT CATTATGAC | 2940 |
| TTTATAGAGC TTATCAAAAT AGATCAGAAC CTTATATATG TTATCTCCTG TAAGCTTTTC | 3000 |
| AGCTTCAATA GTCTGTTTCT TTGCTTTCGC ATCAATTAGT GATGATTCTA ATTCATCTAG | 3060 |

| | | | |
|--|------|--|--|
| 1008 | | | |
| TTTGTGCATAC ATACGATATA GTCTATCATC TAAATCCTGT TTCCTTCTCT TATAATGCTT | 3120 | | |
| ATCTTCAACA TCTAAATTAT CTATTTCCCTC AATTAGCTTA AACTTTGTAG AATGACTCTT | 3180 | | |
| TCTCAATTCC TTTTGGAAT TATCTATTTC TTTTCTATT TCAGAGGTAT CCACCTTCAT | 3240 | | |
| GTGATTTTT TCTTGCATCA TAGAAGCAAA TTTCCGATTA CTTACTATCT TGACAATCAC | 3300 | | |
| CTCTGCAACA GCATCATCTA ACAATTCTTC TCTAATTTGC TTAGTGAATG TACACTTATT | 3360 | | |
| ACCTCTTATC ATCTGCCTAT GGTACAAACC ATAGTAATAA AAATCTTTAT ACTTTGTGCC | 3420 | | |
| ATCTTTCTTT TTCTTGATAC ACTTGTCCC AAACATCCC ACTCCACATA TCGGGCATTT | 3480 | | |
| TACAATTCCA GAAAGCAAGT GTGTGCGTGT ATCTTTTCCT TTATTCACAT GCTCATATTT | 3540 | | |
| CTTTGCTTGA GATTTTAGCT TAACCTGAGC AGCTTGCCAA ACTTCATCGG AAATATAGC | 3600 | | |
| TTCATGTATC CCTTCAGATA TTAGATATTC ATCTTGTTCA ACCTGCTTAT ATTCAATTTCT | 3660 | | |
| TGTACCATGA ACTTTTTCTA AAGTTCTTCT TCCAAATGCT ATTTTCCCAT TATATAcAGG | 3720 | | |
| ATTCTTTAAT ATCTTTCTTA TAAGACCTGC ATCAAACAAA GGATTCTTAC CATCTGTCT | 3780 | | |
| TGGGATTTTT CTAATCCAT GATTCTCTAA GTATTTAGAT ATCCCATGG CTCCTATCGT | 3840 | | |
| AGTATTTACA TACTGGTCGA AAATCGTTCT TATTGCAACT GCCTCTTCT CATTTATAAA | 3900 | | |
| CAGCTTGCCG TCTTCAAGTT TATATCCATA CGGAGCAAAG CCACCATTC ATTTCTCTTC | 3960 | | |
| CCCTGCTTTT TGAATGCGAC CTTCATTGT TTGAATACTG ATGTTTCTC TTTCTATTTT | 4020 | | |
| AGCCACAGCT GATAAAACAG AAATCATTAG TTCCCAGCA TCTTTAGATG AATCAATGCC | 4080 | | |
| ATCTTCAACG CAGATAAGAT TAACTCCATA ATCCTGCATT ATATGAAGTG TAGAAAGAAC | 4140 | | |
| ATCAGCGGCA TTTCTTGCAA ATCTTGATAA CTTAAACACA AGAACAAAAG ATACTCCATC | 4200 | | |
| TTTTCCAGAT TTTATATCTT CCATCATTCG ATTGAAGTGT ATTCTACCTT CAATAGACTT | 4260 | | |
| GTCAGACTTC CCGGCATCTT CATACTCTCC AACAATTTCA TAATCGTTGT AAATAGCAA | 4320 | | |
| AGCTTTCATT CGTGATTTTT GTGCCTCTAA CGAATACCCC TCTATCTGTA TTGACGTAGA | 4380 | | |
| TACTCGTGTA TAGAGGTATA CTTTTATTTT TTCTTTTGAC ATAGTATTAA CCTCAATATA | 4440 | | |
| ATTTTCTAT ATCATATATA ATTTTTTTAA TTTAAGTTTG GACTATCAT TCAAGTATAT | 4500 | | |
| TATAACACTT TTATTAGTCC GTCTCAATTT GTGTTTTTGC CATGTCAAAA CTATTTTCA | 4560 | | |
| TCTCTGATT TTTTGCTGGC GTTGATCGG GTAGATTATC TAAATCTAAA GCACCAGCAT | 4620 | | |
| ATTTTGCAAT CAGATTTGCT ATTAATCAG CCAATCCATT CCAGTCATTG TCCAATATAT | 4680 | | |
| ACCTCCTCTA AAGTTTTATA TCTAATAATT ATTTGTTTAA TTAAGTTTTT TGACATTGAC | 4740 | | |
| AAGTGCTTTG GATTAGCAAC ATAGGAATCT CACTTCCGCC TCTATTCCGG ATGAGCCGGC | 4800 | | |
| TTCAACCTTA GAAGTATCAT TACCCTCATT TTCTTCATAG CGGATAGGGT ATCCCTCCCT | 4860 | | |

1009

| | |
|--|------|
| ATATTCAAAC TCTTACTTAT CGCTCACTTT CTTTTTGCTT AGCAGAACTT TTTTGGCCGA | 4920 |
| ATTATTTCAGC CGAAAGATCT TGACGGATAG GTTATTACGC TCCAAAAATA ATTAACGTCT | 4980 |
| TGTCTTGGTC TATTC AATTG TTAAGGTTCA AAATTTATCG AGAGTTATTA ATCTTTTAA | 5040 |
| AAATTGACCA TCAGAAAATA TTTATCTTGA TGAACAAAA TTCTATAAAT TACCCTCTTA | 5100 |
| TACTTAACAG TGAAGAAG TCTTCTTGG TAACCAATTT TGAAATAGAA TTTGCTTATA | 5160 |
| TAAAAAGGTC CAATTCAC TGCATAAATA GCAGTGAAAA TTAGACCCTC TTGGTAACTG | 5220 |
| TCATCTAAAA GTCTTCTA | 5238 |

(2) INFORMATION FOR SEQ ID NO: 151:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13425 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 151:

| | |
|---|-----|
| GACGATTTAC GAAGAATCGA ACAAGAACCT GCTCCTATCA ATTCCCAACC TCTATCTCTA | 60 |
| AAATCTTGCA GTTCATGCTT ATACTTTTTT AAGAAATCTA GAATCATAGA TACGGTAGAT | 120 |
| GACATCGTCT GGTGACATT GGTCAAAATA GAACAAACCA AACGACTCG TTCTATACCT | 180 |
| CCAACCTTTC AAATGCATCT CATGTAAATG TTCTTCTTCC TTGTCCAAAT CAACAATGGT | 240 |
| GAAAAATCCGA AATCTACTC TGCTATTCAT TGTCTTACCC CAAAATTAGA AAACATGCCT | 300 |
| GGCGTTATTT ATTAGATAAT TCTTCCACT TTTGACTCAA TCTCCAAAAA ATATAAGAAA | 360 |
| TCTGAATCGC AAAAATATC AATAAAACCC AATCTATTAT GAAAATCAAA AACACTTTCC | 420 |
| AACTGAAAGA ACTACCTCCA GTGACAAACT TTGAGAAAAA CGGTAGTAGA GCTAAAAAGA | 480 |
| GAAATAAAAT AGGAAGCATC CGCATTGTTA AAATCCGTTT GGCATAAAAA AATCTTTATT | 540 |
| TAAACGAAAA TATTATGGCA AAATTTACGC CAGTTTTTGA ACGGCTGATG TAGATATTTT | 600 |
| ATACTTTCAA AATGTTTAAA TGTGATTATT TATTTTTGAA AAATAGATCA CCAGCCCGAC | 660 |
| TGAAAGTGCT TATAGAATGA TAATAAGTCG CCTGCCGAAA ACAGCGAAAA ATAGCGGTGT | 720 |
| TATGCGGAGA TAATCTGACG CGATGCGAAA GTATATTGCA TACTTATTTT CAACAATTTA | 780 |
| GCAGAGTATT TTTATAAGTG TGATATAATA GAAGTATAAT TTGTTCTGAT AGTTTATTTT | 840 |
| ATGGAGAAGT AGATTTTAG AATGCGGAGG GTTCAATATG GTTGAGTTTA TAAAGCTAA | 900 |
| GAAAGAAATG AGTGAGGAGG ATATTAAAGC AAATTCATC ACTCCTGCTA TTGTATCCAA | 960 |

1010

| | |
|--|------|
| AGGATGGAAA AATGGTGAGC ATATCGCTTA CGAAGAATAC TTCACTGATG GTCGAATTGA | 1020 |
| AGTTAGAGGA GATAAGGCTC GTCGTAAAGA AGGAAAAAAA TCAGACTATT CACTGTATTA | 1080 |
| CCAATTTGGA ACTCGAATTG CAATTGTTGA GGCAAAGGAT AATAAACACA GCGTTCGAGC | 1140 |
| AGGATTACAA CAAGCTATTG AATATGGAGA GATTTTAGAT GTTCCATTTG TTTATCTTC | 1200 |
| GAATGGTGAT GGCTTTATTG AACACGACCG TATCACGAGA GAAGAACGTG AGCTGGAGTT | 1260 |
| AGACGAATTC CCTACTCGTG AAGAATTATT TTCTCGTATG ACGAAGGAAA AAGGATTGAC | 1320 |
| GTACGAAATT ACAGAAGCTA TCTCAACTCC ATACTATACA GACGCCTTCT CAATGAAAAC | 1380 |
| GCCACGCTAT TATCAGCAA TAGCTATCAA CCGTACTATT GAAACAGTTG CCAGAGGACA | 1440 |
| AAAACGAGTA ATGTTTGTGA TGGCAACAGG AACGGGAAA ACGTTCATGG CTTTTCAAAT | 1500 |
| TATTCATCGC CTTGCAAAAG CTGGTTTGGC TAAACGAGTT TTATTCTTAG CAGATAGAAA | 1560 |
| CATCTTAGTA GACCAAACGA TGGCTGAAGA CTTTAGGCCA TTCGAAAAGG TAATGACGAA | 1620 |
| AATTACACCA AAACCTTTGA CTGCTCCTGA AAAATTAAAT TCTTTTGAAA TTTATCTAGG | 1680 |
| GCTTTATCAG CAACTAACTG GTGAAGATGG AACTGAAACA CATTATCAAA AATTGACAA | 1740 |
| AGACTTCTTT GATTTAATCG TAATTGATGA AGCGCACCGT GGTTCAGCTA AGGAAAACAG | 1800 |
| TAAGTGGCGT AAGGTAATTG ATTATTTTCAG TTCTGCGACA CAGATTGGGA TGACCGCTAC | 1860 |
| TCTTAAAGAA ACCAAGAATG CTTCCAATAC GGAATACTTT GGTGAGCCAA TCTATACTTA | 1920 |
| TAGTTTAAAA CAGGGAATCG AGGATGGTTT TTTGGCTCCA TATCGTGTTA TGAGGGTTAA | 1980 |
| TTTAGATGTG GATGTGGATG GTTATCGTCC AGAACTGGA AAAGTTGATG CTAACGGACA | 2040 |
| ATTAATAGAA GATAGGTACT ACGGCAGGAA AGATTTTGAT AAAACCATTG TCATTGATGA | 2100 |
| TAGAACGCAA AGAGTTGCCA AGTTTGTTTC TGATTATATG AAGCAAAACA ATGCACGATT | 2160 |
| TGATAAAACA ATTGTTTTTT GTGTTGATAT TGACCATGCC GAGCGAATGC GTGCTGCACT | 2220 |
| TGTAAAAGAG AATCTAGACT TAGTCCAAGA AGACTATCGT TATGTCATGC AAGTAACTGG | 2280 |
| TGACAAACGCT GAAGGAAAAG CTCAACTGGA TAACTTTATG GATGTCAATT CTAATTTTC | 2340 |
| CGCTATTGTA ACAACGTCTA AATTATTAAC GACAGGAGTT AATGCTAAAA CATGTCGTTT | 2400 |
| GATTGTTTTA GACTCTAATA TCCAATCCAT GACTGAATTT AAACAAATTA TTGGTCGTGG | 2460 |
| CACACGTCTT TATCCTCAA AGGGGAAAGA ATTTTTTACG ATTATTGATT TTCGAAATGT | 2520 |
| TACCAATTTG TTTGCTGACC CTGATTTTGA TGGTGATCCA GTGAAGGTGC TAGAAACAGG | 2580 |
| TGCGAAAACA GTCAGTGGTT CTACGCCCGG TTTCGTAGAT GAGGAAGGTG ACCCAGTAGA | 2640 |
| AAAATATATC GTTACAGACA AGCAGGTTAC CATTCTTAAT TCTACTGTTC AAGTATTGGA | 2700 |
| TGAAAACGGG AAACGTGATTA CCGAAAGCCT GACCGACTAC ACTCGAAAGA ATATCTTAGG | 2760 |

1011

| | |
|--|------|
| TAGCTACGCC ACTTTGAACG ATTTTATCAC AGTTTGGCAT ACGGCAGATA AGAAGAAGCT | 2820 |
| TATCTTAGAC GAACTTTATA AAAAAGGAGT TTATCTAGAT GCTATTCGAG AGTCGGAGGG | 2880 |
| AATATCAGAA CAAGAAATCG ATGATTTTGA TTTACTCCTA AAAGTGCCT ATGGTCAAAA | 2940 |
| AGAATTAACC AAAACGGAAC GTATCAATAA ACTCAAACAA AGCGGATATT TATATAAATA | 3000 |
| TAGTGAGGAA GCGCGTGCTG TTTTGGAAAT TTTACTGAAC AAATACATGG ATAAAGGTAT | 3060 |
| TGGAGAAGTC GAAAGCATTG AAACATTAAA ACTTCCAGAA TTTTCAGATAT ATGGTGGAAC | 3120 |
| CTTCAAAATC ATCAATACTT ATTTTGGAGA TAAAAACGA TATTTACAAG CAATTAAAGA | 3180 |
| ATTGGAGCAA GAGCTATTTA CAGTAGCTTA ATGAAAGGAA AGTATGTCAA TTACATCATT | 3240 |
| TGTAAAAAGA ATTCAAGATA TCACTCGAAA CGATGCTGGT GTTAATGGTG ATGCTCAACG | 3300 |
| TATTGAGCAA ATGCTTGGT TATTATTCTT AAAAATTTAT GATAGCCGTG AAATGGTTTG | 3360 |
| GGAATTAGAA GAAGACGAGT ATGAGTCAAT TATCCAGAG GAATTAAAAT GCGGAAATTG | 3420 |
| GGCTCATGCT CAAAATGGGG AACGGGTATT GACAGGCGAT GAATTACTTG ATTTTGTCAA | 3480 |
| TAACAAGTTA TTCAAAGAGT TGAAAGAGCT TGAAATAACT TCAAATATGC CTATTCGAAA | 3540 |
| AACGATTGTT AAATCAGCTT TTGAAGATGC GAACAACTAT ATGAAAAATG GCGTCTTGTT | 3600 |
| ACGCCAAGTC ATCAATGTTA TTGATGAAGT TGATTTCAAT AGCCCTGAAG ATCGTCATTC | 3660 |
| GTTTAATGAT ATTTACGAAA AAATCTTAA AGATATTCAA AATGCTGGGA ACTCAGGAGA | 3720 |
| ATTTTATACG CCACGTGCAG CGACTGATTT TATTGCCGAA GTTCTTGACC CAAAAGTTGG | 3780 |
| AGAATCAATG GCAGACCTTG CTTGCGGAAC AGGAGGCTTC TTGACTTCGA CTCTGAACCG | 3840 |
| TTTAAGTAGT CAACGTAAAA CTAGTGAAGA TACCAAAAAA TATAATACAG CTGTTTTTGG | 3900 |
| TATTGAAAAG AAAGCATTTT CTCATCTTTT AGCAGTTACA AATCTGTTTC TTCACGAAAT | 3960 |
| TGATGACCCT AAAATGTTC ATGGAATAC TTTGGAGAAA AATGTTCTGT AATATACGGA | 4020 |
| TGATGAAAAA TTTGACATTA TTATGATGAA TCCACCTTTT GGAGGGTCAG AATTAGAAAC | 4080 |
| AATAAAAAAT AACTTTCCAG CAGAATTACG GAGTTCTGAA ACAGCTGATT TATTTATGGC | 4140 |
| TGTCATTATG TATCGTTTGA AAGAAATGG TCGTGTGGA GTTATTTTAC CTGATGGTTT | 4200 |
| TCTATTTGGT GAAGGTGTAA AAAGTCGCTT GAAACAAAA CTGGTAGATG AGTTCAACTT | 4260 |
| GCATACGATT ATTAGGTTGC CTCATAGTGT CTTTGCACCG TATACAGGAA TCCATACGAA | 4320 |
| CATTCTTTTC TTTGATAAAA CAAAGAAAAC AGAAGAACT TGGTTTATC GTTTAGATAT | 4380 |
| GCCAGATGGT TATAAAAATT TCTCGAAAAC TAAGCCGATG AAGTCAGAAC ACTTCAATCC | 4440 |
| TGTTCTGTGAC TGGTGGGAAA ATCGTGAAGA GATTCTGGAA GGTAAGTTCT ACAAATCTAA | 4500 |

1012

| | |
|--|------|
| ATCATTTACA CCTAGTGAAT TGGCTGAGTT GAATTATAAT TTAGACCAGT GTGACTTTCC | 4560 |
| AAAAGAGGAA GAGGAAATCT TAAATCCCTT TGAGTTGATT CAGAATTATC AAGCGGAAAG | 4620 |
| AGCAACTTTA AATCATAAGA TTGATAATGT ATTAGCTGAT ATTTTGCAGT TGTGGAGGA | 4680 |
| CAAATAATGA CACCAGAACA ACTTAAAGCA AGTATTCTCC AAAGAGCGAT GGAAGGGAAA | 4740 |
| TTAGTGCCGC AAAATCCCAA TGACGAACCT GCAAGTGAAT TATTAAAGAG AATTAAAGCT | 4800 |
| GAAAAAGAAA AACTTATCAG TGAAGGAAAA ATCAAACGAG ATAAAAAGGA AACTGAGATA | 4860 |
| TTTCGTGGTG ATGATGGGAA ACATTATGGG AAGTTTGCTG ATGGAAGCAC TCAAGAAATT | 4920 |
| GATGTTCCCTT ATGATATTCC TGATACTTGG GAGTGGGTGA GGTTTTCTAC ATTGGTTGAA | 4980 |
| ATTGTCAGAG GTGGCTCTCC ACGACCAATC AAGATTATC TTAATTCTGA AGTAGATGGA | 5040 |
| ATAAATTGGA TAAAAATAGG TGATACTGAA AAGGGTGAAA AGTATATAAA TAATGTTAAA | 5100 |
| GAAAAATCA AAAAATCAGG GCTTAACAAA ACTAGATTTG TAAAAAAGG TACATTTTGT | 5160 |
| TTAACTAATT CTATGAGTTT TGGTAGACCT TATATTTTGA ATGTTGATGG TGCAATACAC | 5220 |
| GATGGATGGT TGGCTATTTT GAACTATGAA AACTCATTAA ATAAAGATTA CCTATTCTAT | 5280 |
| ATTCTTTTCAT CAAATGAGT TATTTCTCAA TTTCTATCTC TAATTAGTGG AGCTGTTGTG | 5340 |
| AAAAACTTGA ATAGTGATAA AGTGTCTTCT ATTCTTATCC CTCTCCCCC ACTATCCGAA | 5400 |
| CAACAACGAA TAGTAGAAGC AATCGAATCA GCTTTAGAAA AAGTAGATGA ATATGCTGAA | 5460 |
| AGTTATAATA GACTAGAACA GCTAGATAAA GAATTTCCAG ATAACTAAA AAAATCTATT | 5520 |
| CTTCAATATG CTATGCAAGG AAAATTAGTT GAACAAGACC CAAATGATGA ATCAGTCGAA | 5580 |
| GTTTACTTGT AAAAAATACG AGCAGAAAA CAAAACTCT TTGAAGAAGG CAAGATTAAA | 5640 |
| AAGAAAGATT TGGACATTTT TATTGTTTCC CAAGGAGATG ATAACTCTTA TTATGGGAAT | 5700 |
| ATACCTATGA ATTGGGTGTT TATAAAAAATA AAAGATATTT TTTCAATAAA TACAGGTCTT | 5760 |
| TCTTACAAGA AGGGCGATTT AAGCATTAAT AATAAAGGTG TTAGAATTAT ACGTGGTGGT | 5820 |
| AATATTAAGC CTTTGAATTT TTCTCTGTTG GATAATGATT ACTACATTGA TACACAATTC | 5880 |
| ATCTCCTCTG AGCAAGTTTA TTTAAACAT AATCAGCTAA TAACACCTGT ATCAACCTCT | 5940 |
| TTAGAACATA TTGGAAAGTT TGCAAGAATC GATAAAGACT ATGATGGTGT TGTGGCTGGT | 6000 |
| GGATTTATTT TCCAATTAAC ACCATTGCGA AGTTCAGAGA TTATTTCAAA ATTTCTATTA | 6060 |
| TTTAACTTGT CCTCTCCGTT ATTTTATAAA CAATTGAAAG CAATAACTAA ACTATCAGGT | 6120 |
| CAAGCTTTAT ATAATATTC TAAACTACA CTGAGCGAGC TATTAATTCC GTTAGCTCCT | 6180 |
| TTTGAGGAAC AGGAACTTAT TACTCAAAA GTTGAGAAAC TTTTGA AAA AGTAAATCAA | 6240 |
| CTTTGAAAAT GATTCTTTTC ATCTCTTCAT GATTAGAAAT AGGGATTAAT AATTCGAGA | 6300 |

1013

| | |
|---|------|
| TACTGGTACT ATTTAATGTT TTCCCTTTGA TAGCATCTTT TGAATCACCT AAAGTAGAGA | 6360 |
| TAAGTGGCAA AAATATCATT AAGTAATCTC TGATAATATT TTCTTTATTA GCATAGGGGA | 6420 |
| ATATCGATAT AATGGCTTCA TTATGAGTGG CAGGAATATC CAATATGGCA ACTTTTCCAA | 6480 |
| TAGATAATTT AAAACTCATT AATAAAGTTC CTTTAGGTGA AATGTCTATT TTCTTTGATT | 6540 |
| TTAATGCTAA TTTAGAAATA GATTCTCTCG CATTAGTTAC ATAACCAGAT ATAGGCATAT | 6600 |
| CTGATATAGA TACCCAAGGT ATTCAGTTC CCCAAAAAGT AGCTTCACTG CGTGGAGGAG | 6660 |
| TTTTCCTAT TCTGAAGTTA ACTAGGCTAG CAAATTTAAT ATATCTCCAT GCTTCTGGGA | 6720 |
| TTTCATATAT AGGATAAGAG GTTGTTCGT CTTTGTCCC ATAATAAGAG CCATAATCAC | 6780 |
| AAAAATAGCA GGTAGTCAGT TTGACCACCT GTTATTTTTT ACCAATTAAC AATTTTATCT | 6840 |
| ACAATATTTT GTTGTTCAGT AGCTGTTTC CTTAGATAAA TTCGAGTAGT TTCTATACTT | 6900 |
| TCTGTCCCA TCAAATCTGC AAGCAAGGCA ATATCATTAT ACTTCGCTAA AAAATTCTTA | 6960 |
| GCAAATAAAT GCCTAAAAGA ATGAGGGTAA ATTACGTTAG GATTCATTTT GTATTTATCA | 7020 |
| GCATAATTTT TTAAGTGTG AGCAACTCCT CTTGCTGTAA TTGGTTCGTT AAATTTATTC | 7080 |
| AAAAATAAAT AACCACTTCG GCGATTTTCT GATTCTAACC AACTAAGACA ACTATTTCTT | 7140 |
| AATTTTITAG GAATGTACAG TCTACGAAT TTACCACCTT TTGAGTAAAT GTCAAAATAA | 7200 |
| CCGATTCTTA CATGCTCTAC TTTTAGTTTA ATAAGTTCAC TTACACGAGC CCCAGTTGCA | 7260 |
| CCTAAAAACC AAACGACAAA ATGCCATTTT AAAATACCAT CTTTTTTCAA ACTACGTTTA | 7320 |
| AGAAAAAGGT AATCAGCATG GCTAATGACA TCTTCTAAAA ACGGTTTTTG CTGTACTTTG | 7380 |
| ACAAATTTTA ATTTCAAATC ATCATGACCA ATAAAAGCCA GATATTTATT TACTCCTTGT | 7440 |
| AGTCGCAAAT TGACAGTTT AGGTTTAAAA TTGTCTAATA AATATCCTTT GTATTCAAAT | 7500 |
| AAATCTTCCA TTTTGAGTTC GTAATCTCTC AAGAAAAATC GAACACCATA AAGGTACGAA | 7560 |
| CGCACAGTAT TTTCAGCTAA ACCAGCTTTC TTCAAATGTA ATTCAAAAATC TTCAACGTA | 7620 |
| AAACTCCTAT CTTATGTTTG ATAGAAATTC CACCGCACGT AAAACTATTA TACTAAATTA | 7680 |
| GTGCGTCAAT ATGGGCGAAA AATTGTTTGA TTTTATCAAC GATTCTGGAT TGTTCAAGAA | 7740 |
| GGGTGGGAG GGGGATTAAA TATTCTTTTA TAGTTTTCTG TAATAATTCT TTTTGTTTTG | 7800 |
| TACTACCCGA CGCTTTTCT TCAATAACTG ACTGAACAAT AGGAGAGGAA AGAAAATTAT | 7860 |
| AGATGAAATG GCAATTAATA ACCCCCGATA AGACTCTTAT AACTGTAACA TGGCTATCTG | 7920 |
| CAACAGCCCA GCCATAAGGA TTTTATTTT CATGGTAAAT AGCTAATCGT CCTAACGTAC | 7980 |
| CTAGACCTGT TGAATTCAC ATTAATCAC CATCTCTTAG TAATCTTCT TTCTGGTAAC | 8040 |

1014

| | | | | | | |
|-------------|-------------|------------|-------------|------------|-------------|------|
| TATGAACTGT | TTCGGGATCA | ATAAATCTTG | CTAAGTCAAT | AGAAAAGCCA | GACCATTGAT | 8100 |
| TACATTTCTG | AGCAATCACA | GGGTATATAG | GAATATTTGA | ATATTTTGGA | GACTTCCCTC | 8160 |
| TTTGAATGTA | GGAGGTTATA | TCGTTTAACC | TCACCCATTC | CCAACTTTCT | GGTATTTTAC | 8220 |
| AAGGTACTTC | CTCATAATAA | GAGTTATCAT | CTCCTGGGA | AACAATAGAA | ATGTCCAAAT | 8280 |
| CTTTCTTTTT | AATCTTGCCT | TCTTCAAAGA | GTTTTTGTTT | TTCTGCTCGT | ATTTTTTCAA | 8340 |
| GTAAACTTTC | GACTGATTCA | TCATTTGGGT | CTTGTTC AAC | TAATTTTCCT | TGCATAGCAT | 8400 |
| ATTGAAGAAT | AGATTTTTTT | AGTTTATCTG | GAAATCTTTT | ATCTAGCTGT | TCTAGTCTAT | 8460 |
| TATAACTTTC | AGCATATTCA | TCTACTTTTT | CTAAAGCTGA | TTCGATTGCT | TCTACTATTTC | 8520 |
| GTGTGTTGTC | GGATAGTGGG | GGGAGAGCAA | TTAATAATAG | ATTAAAATTA | TAATCATTGA | 8580 |
| TTGCAGGATA | ACTTGTTCCT | GATAGTTTAT | TATTAACACG | ATTGATAAAA | TTATCTGATA | 8640 |
| ATAAATAATA | TTTCAAATAT | GTTTCGTAA | GTAAGTATC | CAAAACAATA | AATGCTGTAC | 8700 |
| TAGCTATCAA | ATACTCTTTA | AGTTCTCTAA | CTACAGCAAT | ATTTTTTAGA | TATGGTCTAA | 8760 |
| CTGTTGAAAA | TAAGACACTA | TTCTGCGAAA | CTAATTTTCT | AGCACGGGAA | GGCGCTTGTT | 8820 |
| CAGGTGAAAG | ATATGTAGTA | TTTTTGTAGT | TGATTATGTT | CTTTTTTCTA | TCAATACTAG | 8880 |
| ACGTATCTAT | ATACCTAAAG | GATTTCTCTG | GCTTATTTTG | CCCAAATTC | CAATAAATTG | 8940 |
| ATTTTATCCT | CACCCACTCC | CAAGTATCAG | GAATATCATA | AGGAACATCA | ATTTCTTGAG | 9000 |
| TGCTTCCATC | AGCAAATTC | CCATAATGTT | TCTTATGTGC | TTCAAGTATA | TAAAAAGGCG | 9060 |
| TAAAAATACG | CCTATAGATA | ATGGGGTTGA | AATAGGTTTA | TTGTTGATGA | GATTGTAGAT | 9120 |
| AATTCAATTT | TTTACTTCCA | ATCGAATATT | CAATCCTCC | ACCTTTTCTG | CCTGTAATTG | 9180 |
| TTTCATCATA | AATTCAATAT | CTTCAGGATT | TTCCCTTGG | CAACCTCGGC | AGAAATATTC | 9240 |
| TTCCGCTCGA | TCAGGATTCA | AAAATCGACA | AGCACAAACA | AAACAGTCGC | CATCATCATT | 9300 |
| TATTGAGATA | ATATAGTAGA | TTGAAATAAG | ATGTAAACAA | ATCGATTAGG | AAAGTTAAAT | 9360 |
| TAGTTTCTAG | AAATTTT TAG | CAGATGTAGT | GTAATTTCT | AGTCTCAATT | TACTATGGCT | 9420 |
| TCAAAATATAT | CTTTCGAAAA | AATATTTACA | GATGTGTAAT | TTTGAAGCTT | GCAAAAGTTA | 9480 |
| GTAAACTTGT | AGATTTTCGAT | TTGAAGTAAC | TTGTTTCTT | GCCCGATATT | GTTTTTGAAA | 9540 |
| TTGAATTTTT | CCATAGTGAC | TCCTTAATTT | TCTTCTACAC | GTCTGATGAT | AAATCTAATT | 9600 |
| CGCAAAAGAG | TCAAGAGGAT | TTTTCGAAAA | ATAAATAGCG | ACCGAAATCG | CTATTTTAAG | 9660 |
| GGTTATAGGT | ATTTGATGGC | TTAGACTGCT | GTGTGACTGT | TTACCCACAG | GCAATCTTTC | 9720 |
| TTCTATATTA | GTATTAGTAA | AGGTCTAAAT | AATTATCAAT | TTCCCATTGT | GAAACGAAGG | 9780 |
| TTGCATAACT | TGCCCATTCG | ATTGTTTTGG | CTTCAAGGAA | GCTAGTATAG | ATGTGATCTC | 9840 |

1015

| | | | | | | |
|------------|-------------|------------|-------------|------------|-------------|-------|
| CGAGAGCAGC | TTTAACCACT | TCATCTTCTG | TCAAAGCTTT | CAAAGCGTTG | TGAAGAGTTG | 9900 |
| ATGGAAGGTC | TGTAATACCA | GCTTCCTTGC | GCTCTTCTGC | TGTCATGATG | TAGATATTTT | 9960 |
| CTTCGATAGG | AGCTGGTGCT | TCGATTTTAT | TTTCAATACC | ATACAAACCA | ACTTCCAAAA | 10020 |
| GAACAGCCAT | AGCAACGTAA | GGGTTCGCCA | TTGGATCCAC | TGAACGCAAC | TCAAGACGAG | 10080 |
| TTCCCATACC | ACGTGAAGCA | GGTACGCGCA | CAAGTGGCGA | ACGGTTACGA | CCAGCCCAAG | 10140 |
| CAATGTAAAC | AGGCGCTTCA | TAACCTGGAA | CCAAACGTTT | GTATGAGTTA | ACTGTGGGT | 10200 |
| TCATGATGGC | AGTATAGTTG | TAAGCATGCT | TGATCAAACC | GCCTAGGAAA | TGGTAAGCTG | 10260 |
| TTTCTGACAA | CTGCATTCTT | TTTGGATCAT | TTGGATCAAA | GAAGGCGTTA | TTTCCTTCTG | 10320 |
| CATCAAACAA | GGACATATTA | CAGTGCATAC | CTGATCCAGC | AATACCAAAT | TTTGGCTTCG | 10380 |
| CCATAAATGT | TGCGTAAAGT | CCGTGTTTGC | GAGCAATGGT | TTTAACAACA | AGCTTAAAGA | 10440 |
| TTTGAATCTT | ATCACAAGCA | CGGAGAACTT | CATCGTACTT | AAAGTCAATC | TCATGCTGTC | 10500 |
| CAACCGCAAC | CTCGTGGTGA | CTCGCTTCTA | CTTCAAATCC | CATTTTGGTC | AAGACATTCA | 10560 |
| CAATCTCAGC | ACGTGTGTTG | TCCGCAAGGT | CAGTAGGTGC | CAAGTCAAAG | TAGCCACCCT | 10620 |
| TGTCATTCAC | TTCAAGTGTT | GGGTCCCAT | TTTCATCCAA | CTTAAATAGG | AAGAATTCTG | 10680 |
| GCTCTGGACC | AAGGTGAAG | GATTTGAATC | CAACTTCTTC | CATGTGACGA | AGAGCTCGTT | 10740 |
| TCAAATTACC | ACGAGGGTCA | CCCGCAAATG | GTTACCTTC | TGTTGTATAG | ACATCACAGA | 10800 |
| TCAGACCTGC | AACACTTCCA | TTTTCATCTC | CCCAAGGGAA | GACTGTCCAT | GTATCCAAGT | 10860 |
| CCGGGTACAA | GTACATATCC | GACTCATTGA | TACGTACAAA | ACCTTCAATA | GAAGATCCAT | 10920 |
| CAAACATAAC | CTTGTTTCGAC | AAGACCTTAT | CTAACTGTTT | ATCTGTAGCA | GGAATTTTCA | 10980 |
| CGTTTTTCAT | GGTTCCTCAA | ATATCTGAGA | ACATAAGACG | AATAAAGGTA | ACATTTTTTT | 11040 |
| CCTTGACTTC | ACGACGAATA | TCTGCAGCTG | TGATTGGCAT | AAGTTTTCTC | CTTAATCTAT | 11100 |
| GACTACTTGC | GGTTGCCTAA | CCGCGACCAA | AAGGTGACTG | TACTGAAGCA | AAACGCCCCCT | 11160 |
| GTTGGAGGAG | TTCATTGTGA | AGTGCACGAC | GTACTTCAGT | CTGACTAACC | GCTTTCTTGG | 11220 |
| ATPTCGCTTC | ACGTTTCAGCA | TATTTTTTCT | TAATGGCAGC | GATATTATAA | CCTTCAGAGA | 11280 |
| TATAATCTTT | GATTTCAAGC | AGACGATCCA | TGTCATTCAA | GGAATACATG | CGACGATTTT | 11340 |
| CTTCGTTCG | ATCGGGCTTG | ATCAACTCTT | GATCTTCATA | ATAACGAATC | TGACGCGCCG | 11400 |
| ATAGATCGGT | CAACTTCATA | AACTGCGCA | TAGGAAAAAC | AGCCATATTT | CGGCGAAATT | 11460 |
| CTTTTCTCTT | CATTTACAAT | TTCTTCTTCT | CTGTCTATTA | TAGTCTAAAA | AAAGACAAAC | 11520 |
| GTCAATTGAT | AATGTTATAA | AATGTAACAT | TATTTTTTCTT | TTTTCTCTAA | AAAGAGACGA | 11580 |

1016

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| ATACGATCAA | TATCGTAATT | TACGATAATT | GCGACAAAAA | CTCCCATAAA | CGTTTCTAAT | 11640 |
| ACACGCACAA | ACACGTACAA | AATTGTCTCA | CCACTTGGAA | TTGATAGGGT | AATGATTAAAC | 11700 |
| ATAGCTGCTA | CACCACCAAT | AACCCCTGCT | TTGTTATTCA | TGGCTACATT | TGTCATAATG | 11760 |
| GTTAACATGG | TGCAGATTGG | AACAAC TACC | AAGGTCACCC | AAAAGGCTTC | GTGAAAAAAG | 11820 |
| GTATTTAATA | AGAAGAAGAC | CAAGGCATAG | AGTCCACCGA | TACTATTTCC | TAGAATACGC | 11880 |
| GAAGTCCCAA | AATGAACACT | CTCATCAAAA | CTCTCCCTCA | GGCTAAAAAC | GGCTGTCAAA | 11940 |
| GCACCAATTT | GAAGACCTTT | CCAGCCAAAA | AAGCCAAAAA | TCAAGAGAAC | TAGAAAAACA | 12000 |
| GCAATACCTG | TTTTAAAGGT | TCGCATACCA | AGTTTGAAC T | GGGATTTATC | GAATTTATAT | 12060 |
| TTTTTAAAT | AAC TCATAAT | CTCAACTTTC | TATTTCCATT | TTATCATAAA | TCGGTGATTT | 12120 |
| TTATGAGTAA | TAGTTGAGAG | GAAGCGTTTT | TATTTTAAGC | AAAAGAAAAG | AGGAAC TTTC | 12180 |
| ATCCCTCTCT | TCTTTGATTT | ATTTATAAAA | TCTTATTTTT | CTGTCAAGGC | TGCAAGTCCT | 12240 |
| GGAAGAACCT | TACCTTCAAG | AAGTTCCATT | GATGCTCCAC | CACCCG TACT | AATCCATGAG | 12300 |
| AACTTG TCTG | CACGGCCAAG | GTTAATCGCT | GCGGCAGCTG | AGTCACCACC | ACCGATGATT | 12360 |
| GATTTAACTC | CTGGTTGTTT | CACGATAGCG | TCCATCACAC | CGATTGTACC | AGCTTGGAAG | 12420 |
| TCTGGGTTTT | CAAATACACC | CATAGGTCCG | TTCCATACGA | CTGTTT TGGC | ACCAGTCAAA | 12480 |
| GCTTCGTCAA | ATTTGGCGAT | AGATTTTGGA | CCGATGTCAA | GACCAAGGAA | GCCTTCAGAA | 12540 |
| ACTGCTTCAC | CTTCAGTGTC | ACGCACTTCA | GTGTAACCAG | CAAATGCGTT | AGCTTCTTTT | 12600 |
| GAGTCAACTG | GCAAGATCAA | TTTACCATTT | GCTTTT TCAA | GAAGAGCTTT | CGCAACATCC | 12660 |
| AATTTGTCTT | CTTCTACAAG | TGAGTTACCG | ATTCGATAC | CTTGTGCTTT | GTAGAATGTG | 12720 |
| TAAGTCATCC | CACCACCGAT | AAGGACGTTA | TCAGCTTTT | CAAGCAAGTT | TTCGATAACA | 12780 |
| CCGATCTTGT | CTGAAACTTT | TGAACCACCA | AGGATAGCCA | CGAATGGACG | TTCTGGAGTT | 12840 |
| TCAACTGCTT | CTTGGATGTA | GGCAATTTTC | TTTTCAAGAA | GGAAACCAGC | AACTGCTTTT | 12900 |
| TCAACGTTTG | CTGAGATACC | AACGTTAGAT | GCGTGTGCAC | GGTGAGCTGT | ACCGAATGCA | 12960 |
| TCGTTTACGA | AGATACCATC | TCCAAGTGAT | GCCCAGTATT | TACCAAGTTC | AGGATCGTTT | 13020 |
| TTAGATTCTT | TCTTGCCGTC | AACATCTTCG | TAACGAGTGT | TTTCAACCAA | GAGAACTTGT | 13080 |
| CCATCTTCAA | GAGCGTTGAT | TGCCGCTTCT | AATTCAGCAC | CACGAGTGAC | ACCTGGGAAA | 13140 |
| ACAACATCTT | GACCAAGTTT | TGCTGCCAAG | TCAGCTGCTA | CAGGAGCAAG | TGATTTACCA | 13200 |
| GCTTTATCAG | CTTCTTCTTT | CACACGTCCA | AGGTGAGAGA | AAAGAATTGC | ACGTCCACCT | 13260 |
| TGTTTCGATGA | TGTACTTAAT | AGTTGGAAGA | GCTGCTGTGA | TACGGTTATC | GTTAGTGATT | 13320 |
| ACGCCATCTT | TCAATGGTAC | GTTGAAGTCA | ACACGAACGA | GGACTTTTTT | ACCTTTCAAG | 13380 |

1017

TCAACGTCTT TAACAGTAAG TTTTGCCATG TTACAAAAAC TCCGG

13425

(2) INFORMATION FOR SEQ ID NO: 152:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 905 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 152:

| | |
|---|-----|
| GATTTATCCT ACCGnGAAT TTCCGGAGGG GTTCTAGCAG CAATCTTAGG AATCTATGAA | 60 |
| CGAATGATTG GCTTTCTGGC CCATCCCTTT AAAGACTTTA AAGAAAATGT TTTGTAATTT | 120 |
| ATTCCAGTTG CCATCGGTAT GCTTCTGGGA ATCGGCTTAT TTTCTACCC GATTGAATAC | 180 |
| CTGCTTGAAA ATTATCAGGT TTTTGATTA TGGAGCTTG CGGGAGCTAT TATCGGTACA | 240 |
| GTTCTAGCC TCCTCAAAGA ATCAACTCGA GAATCTGACC GAGACAAGAT TGATTAGCT | 300 |
| TGGTTATGGA CAACCTTTAT CATTTCTGGA TTAGGACTCT ATGCCTTAAA TTTGTCTGTT | 360 |
| GGAACCTTAA GCGCCAGCTT TCTTAACCTC GTCCTAGCAG GCGCACTATT GGCCCTTGGC | 420 |
| GTCTTGGTTC CTGGCCTCAG CCCATCAAAT TTACTTTTGA TTTTGGGACT CTATGCTCCT | 480 |
| ATGTTGACTG GTTTTAAAC TTTTGATTTC TTGGGAACCT TCTTTCCGAT TGGAATTGGT | 540 |
| GCAGGTGCAA CTCTCATCGT TTTTCAAAA TTGATAGATT ATGCCTTAAA CAACTACCAC | 600 |
| TCACGCGTCT ATCATTTTAT CATCGGTATC GTCCTATCAA GTACCCTTTT GATCTTAATT | 660 |
| CCAAATGCAG GAAACGCTGA AAGTATCCAA TACACAGGAC TTCACTTGT CGGTTATGTC | 720 |
| ATCATCGCCT TCTTCTTTC GCTGGGAATC TGGCTTGGTA TTTGGATGAG TCAATTGGAG | 780 |
| GATAAATATA AATAATGGCA AAAAAAGTTA AAATCAAAAA AACATTGGTG GAACAAATCC | 840 |
| TATCTAAAGC AGCTATCCCT CATCAGGGGA TTCAAATCAA TGCCCTAGAA GGAGAGCTTC | 900 |
| CTCAA | 905 |

(2) INFORMATION FOR SEQ ID NO: 153:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4278 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 153:

| | | |
|--|------|--|
| 1018 | | |
| CTTGAATTAA ATAAAAACG TCATGCGACT AAGCATTTTA CTGATAAGCT TGTGTATCCC | 60 | |
| AAAGATGTGC GTACGGCTAT CGAAATTGCA ACCTTAGCGC CAAGCGCCCA CAACAGCCAG | 120 | |
| CCTTGGAAT TTGTGGTGGT ACGTGAGAAA AATGCTGAAC TGGCAAAGT AGCTTATGGT | 180 | |
| TCCAATTTTG AACAGGTATC ATCAGCGCCT GTAACCATTG CCTTGTTTAC AGATACGGAC | 240 | |
| TTAGCCAAAC GTGCTCGTAA GATTGCCCCG GTTGGTGGTG CTAATAACTT TTCTGAAGAG | 300 | |
| CAACTTCAAT ATTTTATGAA AAATCTGCCA GCTGAGTTTG CCCGTTACAG TGAGCAACAA | 360 | |
| GTCAGCGACT ACCTAGCTCT CAATGCAGGT TTGGTTGCCA TGAAC TTGGT TCTTGCATTG | 420 | |
| ACAGACCAAG GAATGGTTC TAACATTATT CTGGTTTGGT ACAAATCAA AGTTAATGAA | 480 | |
| GTTTTGGAAA TCGAAGACCG TTTCCGCCA GAACTCTTGA TCACAGTGGG TTATACAGAC | 540 | |
| GAAAAATTGG AACCAAGCTA CCGCTTGCCA GTAGATGAAA TCATCGAGAA AAGATAGAAA | 600 | |
| GAAGAAAAA TGACAGCAAT TGATTTTACA GCAGAAGTAG AAAACGCAA AGAAGACCTC | 660 | |
| TTGGCTGACT TGTTAGCCT TTTGGAAATC AATTCAGAAC GTGATGACAG CAAGGCTGAT | 720 | |
| GCCCAGCATC CATTGGGCC TGGTCCAGTA AAAGCCTTGG AGAAATTCCT TGAATCGCA | 780 | |
| GACCGCGATG GCTACCCAAC TAAGAATGTT GATAACTATG CAGGACATTT TGAGTTGGT | 840 | |
| GATGGAGAAG AAGTTCTCGG AATCTTTGCC CATATGGATG TGGTGCCTGC TGGTAGCGGT | 900 | |
| TGGGACACAG ACCCTTACAC ACCAACTATC AAAGATGGTC GCCTTTATGC GCGCGGGCT | 960 | |
| TCGGACGATA AGGGTCCTAC AACAGCTTGT TACTATGGTT TGAAAATCAT CAAAGAATTG | 1020 | |
| GGTCTTCCAA CTTCTAAGAA AGTTCGCTTC ATCGTTGGAA CAGACGAAGA ATCAGGCTGG | 1080 | |
| GCAGACATGG ACTACTACTT TGAGCACGTA GGACTTGCCA AACCAGATTT CGGTTTCTCA | 1140 | |
| CCAGATGCTG AATTTCCAAT CATCAATGGT GAAAAAGGAA ATATCACGGA ATACCTCCAC | 1200 | |
| TTTGCAGGAG AAAATACAGG TGTGCCCCG CTTCACAGCT TTACAGGTGG TTTACGTGAA | 1260 | |
| AATATGGTAC CAGAAATCAGC AACAGCAGTC GTTTCAGGTG ACTTGGCTGA CTTGCAAGCT | 1320 | |
| AAACTAGATG CCTTTGTTGC AGAACACAAA CTTAGAGGAG AACTCCAAGA AGAAGCTGGC | 1380 | |
| AAATACAAGG TGACGATCAT TGGTAAATCA GCCCAGGTG CTATGCCTGC TTCAGGTGTC | 1440 | |
| AATGGCGCAA CTTACCTTGC CTTCTCCTC AGCCAGTTG GCTTTGCTGG TCCAGCCAAA | 1500 | |
| GACTACCTTG ACATCGCAGG TAAAATCTC TTGAACGATC ATGAGGGTGA AAATCTTAAG | 1560 | |
| ATTGCTCATG TGGATGAAAA GATGGGTGCT CTTTCTATGA ATGCCGCGT CTTCCACTTC | 1620 | |
| GATGAAACAA GTGCTGATAA TACCATTGCC CTCAACATCC GCTATCCAAA AGGAACAAGT | 1680 | |
| CCAGAACAAA TCAAGTCAAT CCTTGAAAAC TTGCCAGTTG TTTCTGTAG CCTGTCTGAA | 1740 | |
| CACGGTCACA CGCCTCACTA TGTGCCAATG GAAGATCCAC TTGTGCAAAC CTTGTTGAAT | 1800 | |

1019

| | |
|---|------|
| ATCTATGAAA AACAACTGG CTTTAAAGGT CATGAACAAG TCATCGGTGG TGGAACCTTT | 1860 |
| GGTCGCTTGC TAGAACGCGG AGTTGCCTAC GGTGCTATGT TCCCAGACTC GATTGATACC | 1920 |
| ATGCACCAAG CCAATGAATT TATCGCCTTG GATGATCTTT TCCGAGCAGC AGCAATTTAT | 1980 |
| GCCGAAGCTA TTTACGAATT GATCAAATAA AACGATAGAA GTCTGAGATC TTATGCTTGG | 2040 |
| ACTTCTTTT GGAGGAAAG TAGATGTCTC AAATCGAAAG AATCAAACAG GCTATCATGG | 2100 |
| CGGATTTCGA GAATGCCAGC TATACAGAGC GTGGCATTGA GCCTCTCTTT GCAGCGCCAA | 2160 |
| AAACTGCTCG CATCAATATC ATCGGTCAGG CTCCGGGACT TAAAACTCAA GAAGCAGGCC | 2220 |
| TTTACTGGAA AGATAAAAGT GGTGACCGCT TCGGGGACTG GCTAGGTGTG GATGAAGATA | 2280 |
| CCTTTTACAA TTCAGGTTAT TTTGCTGTTT TGCCTATGGA TTTCTACTTT CCAGGACATG | 2340 |
| GCAAGTCGGG TGATCTTCCG CCTCGTACAG GTTTTGCAAG AAAATGGCAT CCGCAGGTCT | 2400 |
| TACAGGAATT GCCTGATATT CAGTTAACCC TCTTGATTGG GCAATATGCC CAAGCCTACT | 2460 |
| ATTTACAGGA GAAAATCAGT GGAAGGTAA CGGAGAGGGT GAAACACTAT AAAGACTATC | 2520 |
| TGCCAGCCTA TTTTCCGCTA GTTCACCCAT CACCACGAAA TCAAATCTGG ATGGCCAAAA | 2580 |
| ATCCTTGGTT TGAGGCAGAA GTAGTGCCAG ATTTGAAAAA AAGAATTAAA ACCATTTTAT | 2640 |
| AGTCAATGAA AATCAAAGAG CAACTAGGA AGCTAGTCGT AGGCTGCTCA AAGTACAGCT | 2700 |
| TTGAAGTTGC AGATAAACT GACGAAGTCG GTAACATACG CACGGTAAGG CGACGCTGAC | 2760 |
| GTGGTTTGAA GAGATTTTCG AAGAGTATTA GAAGAAAAAG AATGAAAGAA ATAGCCTTTG | 2820 |
| ACGCATTTTA CCAGCTTTAC CAAAACGACC AGCTTCTTTT AGTGGATGTG AGAGAAGTGG | 2880 |
| ATGAGTTTGC AGCTCTTCAT TTAGAAGGTG CCCACAACCT ACCGCTTAGT CAATTGGCTG | 2940 |
| ATAGTTATGA TTAATTGGAC AAAGATCGCT TGCATTATAT TATTTGCAAA TCTGGAATGA | 3000 |
| GATCGGCGCG TGCTTGCCAA TTCCTATTAG AACAAGGTTA TAATGTTATC AATGTCCAGG | 3060 |
| GTGGCATGTT AGCCTTTGAA GAACTTTAAA ATTTTGCAAT TCTCCTACTT GGTGTGGACT | 3120 |
| GGGTAGGAGA GTTTTATTTT TAGATAATTC TTATTTTAA GAAAATTGAA AACATTTAAT | 3180 |
| ATTTGCCTCG TGATGCTTTT TTCAGACTCC TAATCGTGGT ATACTAGGTC AGTATTTTAT | 3240 |
| AAATATGAAG GAGATTTTGA TGGCTAAAAA AGGTACCCTA ACAGGTTTGC TCCTGTTTGG | 3300 |
| AATATTTTTT GGTGCGGGGA ACTTGATTTT TCCGCCTTCT CTAGGTGCTC TATCTGGAGA | 3360 |
| ACATTTTCTT CCTGCCATCG CAGGTTTGT CTTTTCAGGC GTTGGTATCG CCGTCTTGAC | 3420 |
| CCTTATTATT GGAACGCTAA ATCCTAAAGG ATATATCTAC GAGATTTCAA CGAAGATAGC | 3480 |
| GCCTTGGTTT GCGACTCTTT ACCTCTCAGT TCTTTACTTG TCAATCGGTC CATCTTTGCG | 3540 |

1020

| | |
|--|------|
| TACCCACGT ACTGCTACAA CAGCTTACGA AGTAGGGATT AGCCCCCTTT TGTCCGATGC | 3600 |
| AAATAAAGGA CTGGCTTGA TTGTATTTAC GGTTCGTAT TTTGCGGCAG CCTATTTGAT | 3660 |
| TTGCTTAAT CCATCAAAAA TCTTAGACCG CATTGGACGT ATTTTAACGC CAGTCTTTGC | 3720 |
| AATTTTGATT GTTATCTTGG TCGTCTGGG AGCTATCAAA TATGGTGGAA CAAGTCCTCA | 3780 |
| AGCTGCTTCA CTGCTTATCA AGCTTCTGCC TTTGGTACAG GTTTCCTAGA AGGTTACAAT | 3840 |
| ACCTTGGACG CCCTTGCCTC AGTGGCCTTT AGCGTAATCG CAGTTCAAAC CTTGAAACAA | 3900 |
| CTTGGATTTT CAAGTAAGAA AGAATACATT TCAACTATTT GGGTTGTTGG TATCGTTGTT | 3960 |
| GCCCTTGCCT TCAGCGCTCT TTACATCGGT TTAGGTTTTT TGGAAATCA TTTCCAGTA | 4020 |
| CCAGCTGAAG CGATGAAGGG TGGAACACCA GGTGTTTACA TCTGTGTCACA AGCCACTCAA | 4080 |
| GAAATCTTTG GCTCAACAGC TCAACTCTTC CTGTCAGCTA TGGTTACCGT AACCTGCTTC | 4140 |
| ACAACGACTG TTGGTTTGAT TGTGTCAACA GCTGAGTTCT TTAATGAGCG CTTCCCACAA | 4200 |
| ATCAGCTACA AGGTTTATGC GACAGCCTTT ACCTTGATTG GATTTGCTAT TGCCAATTTG | 4260 |
| GGTCTTGATG CGATTATC | 4278 |

(2) INFORMATION FOR SEQ ID NO: 154:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1953 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 154:

| | |
|---|-----|
| ACCCGATCAA ATGACAAAAG CTAACCTTGG TGTCGTAGGT ATGGCCGTAA TGGGTCGTAA | 60 |
| CCTTGCCCTT AATATTGAAT CTCGTGTTA CACAGTTGCT ATCTACAACC GTAGTAAAGA | 120 |
| AAAAACGGAA GATGTGATTG CTGCCATCC TGAAAAGAAC TTTGTACCAA GCTATGACGT | 180 |
| TGAAAGTTTT GTAAACTCAA TCGAAAAACC TCGTCGTATC ATGCTGATGG TTCAAGCTGG | 240 |
| ACCTGGTACA GATGCTACTA TCCAAGCCCT TCTTCCACAC CTGACAAGG GTGATATCTT | 300 |
| GATTGACGGA GGAAATACTT TCTACAAAGA TACCATCCGT CGTAATGAAG AATTGGCAAA | 360 |
| CTCTGGTATC AACTTTATCG GTACTGGGGT TTCTGGTGGT GAAAAAGGTG CCCTTGAAGG | 420 |
| TCCTTCTATC ATGCCTGGTG GACAAAAAGA AGCCTACGAA TTGGTTGCGG ATGTTCTTGA | 480 |
| AGAAATCTCA GCTAAAGCAC CAGAAGATGG CAAACCATGT GTGACTTACA TCGGTCCTGA | 540 |
| TGGAGCTGGT CACTATGTGA AAATGGTTCA CAATGGTATT GAGTACGGTG ATATGCAATT | 600 |
| GATCGCAGAA AGCTATGACT TGATGCAACA CTTGCTAGGC CTTTCTGCAG AAGATATGGC | 660 |

1021

TGAAATCTTT ACTGAGTGGG ACAAGGGTGA ATTAGACAGC TACTTGATTG AAATCACAGC 720
 TGATATCTTG AGCCGTAAAG ACGATGAAGG CCAAGATGGA CCAATCGTAG ACTACATCCT 780
 TGATGCTGCA GGTAACAAGG GAACTGGTAA ATGGACTAGC CAATCATCTC TTGACCTTGG 840
 TGTACCATTG TCACTGATTA CTGAGTCAGT GTTTGCACGC TACATTTCAA CTTACAAAGA 900
 AGAACGTGTA CATGCTAGCA AGGTGCTTCC AAAACCAGCT GCCTTCAACT TTGAAGGAGA 960
 CAAGGCTGAA TTGATTGAAA AGATCCGTCA AGCCCTTTAC TTCTCAAAA TCATTTTATA 1020
 CGCACAAGGA TTTGCTCAAT TGCCTGTAGC CTCTAAAGAA AACAACTGGA ACTTGCCATT 1080
 TGCAGATATC GCATCTATCT GGCCTGATGG CTGTATCATC CGTTCTCGTT TCTTGCAAAA 1140
 GATTACAGAT GCTTACAACC GCGATGCAGA TCTTGCCAAC CTTCTTTTGG ACGAGTACTT 1200
 CTTGGATGTT ACTGCTAAGT ACCAACAAGC AGTACGTGAT ATCGTAGCTC TTGCGGTTCA 1260
 AGCAGGTGTG CCAGTGCCAA CTTTCTCAGC AGCTATTACT TACTTTGATA GCTACCGTTC 1320
 AGCTGACCTT CCAGCTAACT TGATCCAAGC ACAACGTGAC TACTTTGGTG CTCACACTTA 1380
 CCAACGTAAA GACAAAGAAG GAACCTTCCA CTACTCTTGG TATGACGAAA AATAAGTAGG 1440
 TCAGCCATGG GGAAACGGAT TTTATTACTT GAGAAAGAAC GAAATCTAGC TCATTTTTTA 1500
 AGTTTGGAAC TCCAGAAAGA GCAGTATCGG GTTGATCTGG TAGAGGAGGG GCAAAAAGCC 1560
 CTCTCCATGG CTCTTCAGAC AGACTATGAT TTGATGTTAT TGAACGTAA TCTGGGAGAT 1620
 ATGATGGCTC AGGATTTTGC AGAAAAATTG AGCCGAACTA AACCTGCCTC AGTCATCATG 1680
 ATTTTAGATC ATTGGGAAGA CTTGCAAGAA GAGCTGGAAG TTGTTACGCG TTTTGCAGTT 1740
 TCATACATCT ATAAGCCAGT CCTTATCGAA AATCTGGTAG CGCGTATTTT GCGCATCTTC 1800
 CGAGGTCGGG ACTTCATTGA TCAACACTGC AGTCTGATGA AAGTTCCAAG GACCTACCGC 1860
 AATCTTAGGA TAGATGTTGA ACATCACACG GTTTATCGTG GTGAAGAGAT GATTGCTCTG 1920
 ACACGCCGTG AGTATGACCT TTTGGCGACA CGG 1953

(2) INFORMATION FOR SEQ ID NO: 155:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 6474 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 155:

CCGGCAGTAC ACGAGCTTGG GGAACAGCCA CTGGAACGAT GAGGTGTGAG CTCAAAATAT 60

| | | | |
|--|------|--|--|
| 1022 | | | |
| CCTCCAGTTA TGTMTTCCT AATAGTATAC CGGAAGAGTG AAAGGATTTT ATAATGGAGC | 120 | | |
| GGTTACAAAG AACCTACTTT CTATTAAACA GTATACTATG AAAATGTGAA AATTTAACAT | 180 | | |
| TTTTTTGTAC AAATTTTATA AATTATTGCC TTTTAAATAT CAATAGTTAA TCTCTTATCC | 240 | | |
| AGATCCCCCT TGTGTAAACT TTATCTTTAT AAGCTTCAAG GCCCTATCC CATCTATTTG | 300 | | |
| CAACAATTAG ATCACTTTGT TTTGTAAATA GTTCAAAATT CTTTTCATA ATTACGTTAT | 360 | | |
| CTATACTAAC GTTTAAATTT GGTTCATATA CTAAAATTTT TATACCGACA ATCAATAGTT | 420 | | |
| CATTAATTAT ACTTAAATA GCTGACTCTT TGTAATTATC TGAATTATAT TTCATCCCCA | 480 | | |
| ATTTATATAT TCCTACTATC TTTGGCTTTC GTTCCAATAT TTGTTTAACT ATGAACTGTT | 540 | | |
| TTCTATTTGT GTTTGAAATA TCAATCGCTT CTATCACTGG GGCATTTATT TCTATAAATT | 600 | | |
| CTTTTTTTTAA TTGTTTAGTA TCTTTGGGAA GACAATATCC TCCAAATCCA AAAGAAGGAT | 660 | | |
| TATTATAAAA ATTTCCAATT CTTGGATCTA AACAAACACC TTTTATTACA ACTTCAGCAT | 720 | | |
| TTAAGCTTCT CCTCTCAGCA AAAGAATCTA GTTCATTAAA AAAGCAACAC GGAGAGCTAA | 780 | | |
| GAATGTGTTA GAAAAAGCT TAATTGCTTC TGCTTCAGTA GGAGAACTA ACATAACATT | 840 | | |
| TTTAATATTG GCAGTACTAT GAGTACTAAT CGAAAGGAAC AACTCTGCAA TTTTCTTCC | 900 | | |
| TTCAACTGTC TCATCTCCAA CAACTATGCG ACTTGGATAT AAATTATCAT ATATAGAACA | 960 | | |
| ACCTTCTCTC AAAAATTCAG GGACAAAAAT GATATTTTTT GTATCAAACA GCCTTTTTTAA | 1020 | | |
| TTTGTTTGAA AAGCCGATCG GAACTGTGA CTTTAAATA ATCTTTCCAT TAGGTTTTAC | 1080 | | |
| CCTCAGAATC TTCGATACCG TTTGTTGAT TTCATATGTA TTAATACTAC CAATTTCTC | 1140 | | |
| ATCATAATCT GTCGGAAGCG CAATAATATA ATAATCAATA TTATTTTAA TTTCAGAAAA | 1200 | | |
| TGTATCAAAA AAAGTAATAT TTAAGTTATT CTCGCAAAAA AACTTCATAA GCTCTTCATT | 1260 | | |
| TTTAGATGGA AGAATGCCCT TTTTAAAT ATTTATTTTT ACAGAACTA TATCATATGC | 1320 | | |
| AACAACCTTA TATTTAGATG CAAATAGTAA CGCGTAGGCC AGCCCAACAT GCCCAAACC | 1380 | | |
| AATTACTGCT ATATTCATAA AACTACTTCC TTATTTCTTA ATCCAAAATC TAATAGAA/A | 1440 | | |
| AGCTGCCCCA TTCCTTAAAT ACAACTCTTT AATATTGTTT AAAAGTTTTT CAACTGATTT | 1500 | | |
| CCAGATTATC AAAATCTGAG ATTTATAGCA CAATATTGAT GATATTCTAT CAATATAATT | 1560 | | |
| TTTTTCATCA AGTTCCTCTT GATACATTTT TAATCTTTA GTTTTTCCCA TATAACTAAC | 1620 | | |
| CATACTACTA TCACTTACAT ATGGGAAGTC CTCATAATAT ATTACTTTAT AACGCATAAA | 1680 | | |
| TTCAAGCGCC CTTCCAATAC TATTCACAAA AACATGAGCA ACATGGTCAC CAAGTGAAAG | 1740 | | |
| CGGACAATAT ACGACACATT TGTCGTCTAA ATGCATTAAC AGCTCTTTTA TGATATCATT | 1800 | | |
| CTTTAATGTG TCCTCATTTT TTAATTCACT ATAGATATGA CGGTATAGAA AATTGCCATT | 1860 | | |

1023

| | |
|--|------|
| TCTATCTTTC CTATAGAGAC ATTCATAGTA CGATAAGTGT CTAAAAATCAC ATTGTAGACG | 1920 |
| TTCACAAGCT AACCTGTCTT CTTCTTCCTT TTCTTCAATC GGATATTTC CAAGGTTACA | 1980 |
| CAACTTATGA AATTGCTTAG CAGAGGGCTG TAGCTGTTGG CTCAAAGGGT AACCAGAAAA | 2040 |
| TATAGTAATA ACAAGTACAA TTTCTCCTTC TGAAGTTAAT TTTGAAATAT AATCACCACA | 2100 |
| GGAAAAAATT GCGTCATCTA AATGTGGAGA TAAAAAGATA TACTTAGTAT TGTTACTCAT | 2160 |
| AACCATTCCC TCTACAATT ATCTAAAAAC TCACTAAGTG TCTGATTAAA TTCCACATCA | 2220 |
| TCAAAAAAAT TCACCTTATT CTTAATAATG AATATTTCTG TAAATAAACA TATATATAAA | 2280 |
| TATTTCAATA TCCTTTCAAT ATCATCCTCT AAATTTCTCT CAATATTTTG TATCAGCCCA | 2340 |
| TTTACAATCT TATTAAAAA GATAAGCTCT TTATCTCTAA AATTAAATAT TTTCATACAA | 2400 |
| CTGTTGTATC GAAAAATATA TAAAAAATT TTTACTAATG TTTGAATATT TAAACAATA | 2460 |
| AATAAATGAG TTGTACCCGG GACACTATTT ATGTTATCAA GAACACTATC TTGAAACCTC | 2520 |
| AACTCACAGT TCTTTTGTG AAATCTTTT TTATCGTTA GATCTGATAT TTTTTTAGAC | 2580 |
| ATTTCAACAA TCTCAGACAT TTTATATGGA TATCTAGGAT GAATGCCAAA ACTATGCAA | 2640 |
| ATGAACTGCA CCCCCAAGT TAGACAGAAT AAATCTAACT TTTGGGGTGC AGTTCATAAG | 2700 |
| ATTGGGATAT TTTTTTGTAG CTAGAACTAG TAGAAATATA TAGTCAAATA ACAGATACCT | 2760 |
| TAAGGGTTTC TCATCTACAT AAAAAATGA TACTTTTTC TCTTCAGTAA TTACCTCATA | 2820 |
| AGCTTCACAA TAGAATCTCA TGTTTCCCTC CCCTATATTC TTAAATAAAA TCCTTTGGAA | 2880 |
| ATTGATATAT CTTAGTAAAA TATTGTTTAA GTTCCGGATG CGGAGCATGG GTAACAATA | 2940 |
| TGACAGTCAA ATCCTCTCTA TCTAATATCT TACGTTCAAT CGCTAACGAA GTTCTCCTAT | 3000 |
| CGATAGCAGA AGTTCCCTCG TCAATTAATA CTATTTTCTT ATTTCTAATT AGCCCTCTAG | 3060 |
| CTAAAGTAAT TTTTGTGTTG TGCCCTCCTG ACAGTAATCT CCCATCATCA CCAACATAAT | 3120 |
| AATCTAAAAT GTTATTAGGA AAATCTTTTA CACTCAAACC AACTTGCTCT AAAGACTGTA | 3180 |
| GTATTTCTTC ATCAGTATAA TTTTCTTCCA ATAAAAATAT ATCTCTAATC GTACCTTCAA | 3240 |
| ACAAATAAGC TTTTGTATCT ACATATAGAA CATTCGAAAC CATATTTAAA TAGGAGGTTT | 3300 |
| TTTTTATATC ATCCCCGAG AATCGCAATT CTCCACTATA ATCTCTCAA AAGCCATTCA | 3360 |
| ATAATTTTAA TAATGTAGAT TTCCCGCTTC CACTTTCACC TAAAATTAAA TACTTTTCAT | 3420 |
| TACGTTGAAA ACAAAAAATT AAGTTTTTAA ATATTTCTTT ATCTCCATAC TTATAGCAAA | 3480 |
| TATTTTTTGC TTCATATAAC GGAAATCTC TATTCACCTC ATTTGGTTCG ATATCATTCA | 3540 |
| TTTTATTTGA CTCAATTGGA TTAATTGAAT ACAATTTTAA AAAAATAGGC TTCGTACCAA | 3600 |

1024

| | |
|--|------|
| TAATAGAGGA TAATTGACCT CCTAATTCAC CTAGCGCTGT AAAAATAACA CCTGTTAGTG | 3660 |
| CTCCTATTGC TTCAATAGTA CCAATTTTCA CTATTCCTTT TATTGCAAGA TAGCCTGTTA | 3720 |
| AAAAAACGAG AGATATCTGA AAAAAATAT TGAGAAAGAA GCTAATAGCG CCTGCTAACG | 3780 |
| TTTCTACAGT TGTCTTCTT TGTATAACCA TCTTTAATAA AATTCCTGCT TCTTTAATTT | 3840 |
| TCTTAGGCAA TACATATAAA AGATTCAAGG ACGCTAACAC ATCAAATCCA TTCAATATAG | 3900 |
| TCTCACTAGA TTTTAAAAA GCTTCATTTT GGTAGTTAA ATTTAGACTA ACTTCTCGCA | 3960 |
| TTTTCGATGC AAAGATTTT GGTACAAGTA GCATAATCAT TAATGAAAAC AAGGTGGCTA | 4020 |
| CAGTCAATGA CCAATGATAG TGATTAAGAG TCACAACGCT AAAATATAGTA CCAGAAATTC | 4080 |
| CTTTTATTAC TAAAAAAGT TGTTTAAACG CCTGATCATT TAAAGTCTGA ACATCATTAT | 4140 |
| TTAGCCACGA AAGATATGTT CCTGATGATT TACTATGAAA TTCTTGATAG GTAGAGTTAG | 4200 |
| AGATGTCTGT GGCAACTCTA TTTCGAATCT CTAGATTAAA CTCTTGGATC ACTTCAACCT | 4260 |
| GATAATTTT CACTACCCAG TCAAGGAATA TTATCCACA CCAGACAATC ATTTGGTAGA | 4320 |
| TTGACAATTT CAAAAACGCT TCTAAATCA TCGCAATTAA TTCATTCAAC ACCAGAGCAT | 4380 |
| TAATAGTTCG TGCATAAATT AGCAATAATT GACCAGCAAC AATAAATATC GTTAATAAAC | 4440 |
| TAAATTTTT TATATTTGAT TTTATAATAG TATACACAAT AGTTTCTCAC TTTCTAAATT | 4500 |
| TTAATTGAAC ATAGTTTCA TATATACAAT AGAAAAACC AAAATGATAT AATAACATAT | 4560 |
| ATTTCAAAAA AGAAATTCGT TAAAAATTT TTCTTCTCTT GCCTTCTTGA TTACTTTTAA | 4620 |
| AGCCTTGCAAT TTGTCTCCTA TTAATAGTAA CCGCTTTATG TTAAAGAAT AATATTTCTT | 4680 |
| TGTAACCAAT ATTCTCTCGT TGAACTCAA TAAATTAAAA TATTTCTTAC AGTAATTATA | 4740 |
| ATATTCCTCA TCTGCATTAA TTGTTTTTTG TGTCACTCCA GTGATACCGT TTTCTTTACT | 4800 |
| GTGAGCGTAG TAATTCACCA AGAATTCCTG CACTATATCA ATTTGGTATC CTTGAACAAG | 4860 |
| TAGTTTTAAT AAAACAACAC CGTCCTGATG TGAATCTATT TTCTCAAAAC CATTAATTAA | 4920 |
| TTCTAGCACC TCTTTTTTAC ACAACCAAAA TGACGTACCT GCTATATTGT GAACCATTG | 4980 |
| AACAAACAAG GGATTTCCAA CAAAATCGGT CTTCCTCTCT TCTCGTGATC CATTTGGATA | 5040 |
| AATTATTATT CCATAACTAC AACTAAAGC TAAATCTTTC ATTCCTACTCT TTTTAAACA | 5100 |
| AGCCATCAAC TTTAAATTC GATCTGGCAT ATATTCATCA TCATCGTCTA AAAATGATAT | 5160 |
| ATACTTACCT CTAGAATTTT TGATACCTAT GTTCTGGCA TTAGTTGCAC CTAAATCTTC | 5220 |
| ATTAATTAAA ATTAATTAA TTCTATGATT GGTATAGCCA AATTGATGGA TAATTTTATT | 5280 |
| TCTTAAATTT ACATTACTAT AATTATCATC AATAATTATA ACTTCGATAT TTTTATAACT | 5340 |
| TTGATGTAAA CAACTTTTCA CAGCTCTAAT CAGAGATTCA TACCTATTAT GTGTTGGTAT | 5400 |

1025

| | |
|---|------|
| TATAATACTT ACTAATTCTT GATCTATATT CCTATCCATG ACTACTCTTC TCTAATAATT | 5460 |
| CATCATATAC TCTCATGGTT TCTACAAACA TTTTGTGCAC AGAAAAATGT TTTCTTATTT | 5520 |
| TTGATTTACT ATTCTCACCT ATATATTTCA AATACTCAGA ATCATTGAGT AAAAAATTAG | 5580 |
| CACAAGCACA CACTCCCTCA ACATCTTCCT TCTCAAATAA AAATCCATCA ACCCTATGTT | 5640 |
| CAATAATTTC ACTTAACCCG CCAACATTAC TAGCTAAAAC CGGAGTTCCT TGTGACATTG | 5700 |
| ACTCTAAAAC ACACATAGGT ATTCTTCTG TATCAGAAGG AATATACAAT AAATCCGATA | 5760 |
| TTTGGTAAAC TATAGTAGCT GGATAGATTT CACCAAGTAA CCTGAAATTA TCTCTACATT | 5820 |
| TCAAATGGCA AATTTTCTT TCAAAAGCAG CCCACATACT ACCATTTCCTA GCCATAATAA | 5880 |
| AAATCACATC TTCTCTGACT AAAATAAATT TTTCTGCAA TTCAAGGAAT CTATCCGGCC | 5940 |
| TTTTTCTGG ATCCAACCTT CCAACATAAC AAATGATTTT TTGTTATTTG GAATACAAAA | 6000 |
| TTCTTTTTTA AAGTCTTGAA CACCTACTAC ATCTAAATCG CTATTTGATA CATTAATTCC | 6060 |
| GTTATTTATT GCAACTATCT TCTTATTTT TATTATACTC TCCAATCTTT TTTTTCATAG | 6120 |
| TTTCAGATAC ACAAATAAAA GCATCTCCCA TAGAATATGT CCAAAAATCA AAATAAGTCA | 6180 |
| AGAATTTCTT TTTTAAGTTA TATTCAACCC ATCCATGGCA TGTTATCACT GTCTTAACCT | 6240 |
| TTCCAAATCC ATTCTTGTC AATTTTTTTA ACATATATAA AAAATAATTA GTTGAGTAGC | 6300 |
| CATGACAGTG TATAAGTTGG ATTTTAAATA ATTTTAAAT ATTTTAAACG TGTAAGGCAG | 6360 |
| TTTCAAAATT ATTTGAACAT TGAGTACAAT CAACATAGGC AATATCTAAA TTTTATAAT | 6420 |
| CATCAATAAC CTTGAATCT CTAGATACAA TTATCAAAAT AGGGAATAGA GACA | 6474 |

(2) INFORMATION FOR SEQ ID NO: 156:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 4792 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 156:

| | |
|--|-----|
| TATTTAACGA TTTTTCAT GTCATTCCT CAAAATAGA ATACCTTATA ATCTAACAG | 60 |
| AAAAAGAGCA TTTACGCCAT TATATGATAT CTATCTCTGT GATAAGTTT TTTTATGGGT | 120 |
| AATTTAAAAG ACCAAACGCA AGATGGCAAT CAAGACCACT CCAAAGAGAA CTGTTCCGAC | 180 |
| TAGATTGCGG TAGCGAAAGG CTACCAAGC TGTGGAAAG ACGGCTAAGA AGTCCAGTCA | 240 |
| TTTGATTTGA GGAAGACTGC CAACCTTACC TGTCACCTACG CTTGAAAGAA TCAGGGCAAA | 300 |

1026

| | |
|--|------|
| GATAATGGAA ACAGGCAAAA ACTTCAAAAA ACGCTCAACA ATCGCAGGCA GCCCCTTATA | 360 |
| CTTGACCAAG ATGAAGGGAA TCATACGGGG AATCCAAGTC ACCAAGCCAG AGAAAAATAAC | 420 |
| TGCTAATAAA AGATACTTAC TGACCATCTA AAACCACCCC CATGCTACAA CCAAGTAGCG | 480 |
| TCGCAACAG AACAGCTAGT GACTGAGACA TCACTGTCAA GAGCAAAAAG AAGGACACCG | 540 |
| CAACAACCTGC TAGGATAATG AGCAGATTGC GGACAGGAAT CCGTCTTTGC ATAATCTGAA | 600 |
| ATTGCGAAGC AAAATACCAA TAAACATCCC AACCAGGGCA AAATCCAAGC CAAAGATTTT | 660 |
| TGGATTGGT AGCAGGCCAC CCAGAGCCGT TCCGACTACT GTCCCCACAA ACCAAGCCAC | 720 |
| ATAGCTGTTA AGATTGTTTC CGTGCATCCA CATAGGATTT ACCTTGTCTG TATGGGCCAA | 780 |
| TTCAACCATC AAAACGCCAT AGGTCTCATC TGTCAAGATA CTAGACATAC CGATATTGTA | 840 |
| CCAAAGACTG GTATGACGGA AATAAGTCGA TGCGTGTAAG CTCAACAAAA AGAGACGCAA | 900 |
| GTTGATTAGA AAAACCGTCA TAGCAATAGC TGCCACAGGA GCTTGAACCA CAATCAGTGC | 960 |
| CAACATGGCA AACTGGGCAC TCCCAGCATA AACAAAGAGA CTCATCAAGC CCATCTCAAC | 1020 |
| AGGTGTCACA TAGGGCGCAC CGATAATTCC ACAGGCCAGG CCGATACTGA CATAGCCAAG | 1080 |
| AGCCGTGGC ATGGCTGCCT GCGCCCCCTC CTAAATCCT TTTTCTTTCA TCTTTCTCCT | 1140 |
| CATATTGTCT TAATAATACT CAATGAAAA CAAAGAGCAA ACTAGGAAAC TAGCCGCAGG | 1200 |
| TTGCTCAAAA CACTGTTTTG AGGTTGCAGA TAGAACTGAT GAAGTCAGCT CAAAACACTG | 1260 |
| TTTTGAGGTT GTGGATAGAA CTGACGAAGT CAGCTCAAAA CACCGTTTTG AGGTTGTGGA | 1320 |
| TAGAACTGAC GAAGTCAGTA ACCATACCTA CGGCAAAGTG AAGCTGACGT GGTTTGAAGA | 1380 |
| GAGTTTCGAA GAGTACAAGT AGGCTGAAAA GAATCCAACC ACAGCATGGA CTATTATATA | 1440 |
| GCAGATTGAA ATAAGATGAG AACAAATCGA TTGGGAAAGT AAAATTAATT TCTATAAATG | 1500 |
| TTTTAGCAAT TGTTCGTAC TATTTTAGAT TCACTCTATT ATAACACATT CAGAAAAGAG | 1560 |
| AAAAAGTCT GTTGATTTTG ACCATCATAA AAAGACTGGC AATCCAGTCT CAAACATATA | 1620 |
| TTATAGAAAT TCTCCACTAA ATACTTTCAC GAATATTCAG AAGCATAACA AAGGCAACTA | 1680 |
| GAAGAAATAG CAATAAAACA AAGCTAACTG CCAGAGTTCC AAAGCTAGTA GCAATGGTTA | 1740 |
| CCAAAGCTAT TGTAAATAAG CTAGGTAAAA CAACCGTAAT GGCACCGATA GAGGATTGAA | 1800 |
| CTGCTCCCAT TGACTCCTCA GGTATTTGTT TAAAAACGAG TTCTTGCAAT CTAGGAGAGA | 1860 |
| GAACACCTGC GAAAAAGGCA TCCAAGGTAC TAAAGATGAG AATCCAGTCA AAACGAACTG | 1920 |
| TGGCAAATCC TACTAGAAGA AGCAACTGGA TGACAAGTGA GGCATAGAGA GCTGTTTTTA | 1980 |
| TGGAAATGGT ATGTTGCAGA TAGCCACTTA CAAGGCTTCC GACAATCAGG GCTGATAATT | 2040 |
| CTAGTGTGGC TAACAAGGCA AGAGATTGAC CAGTTTGTA ATTCAAAAAG GGCTGGTTCC | 2100 |

1027

| | |
|--|------|
| TTAAAAATAG AGTGGAATA GGAACCGTAA CATTTATCAC TGCTTGACTA GTAGAGATAA | 2160 |
| TAAACAAAAC CAAGAGCACC TTATTCATAT TCCATATCAA TTTCGATGAT TGGAGCAAAT | 2220 |
| GCTGGCAAAA GGATTTTACA GAGAGTCCTT CTTGATAGCT AATCGTTTTT TCTACTTTCA | 2280 |
| AGAGGTCACT TTTTATGAAG AGGATACCTA AAAATGCGAT TAAAAAGGTA AGAGCGTTCA | 2340 |
| GTAAGGAAAT AAACCTGGATG GATAGAATGC CTAGTAAGAC TCCTCCTAGG ATATTACTGA | 2400 |
| TTGTTTTTAC TAAACTAACA GTTGACTGTT TAAAGCCAAT AGCTTCTGCC AGATGGTCTT | 2460 |
| GCCCAATAAT TCTAATGAAA ATCGGAGTGA GCATGGCGCC TGAAAAATAA CTCAATGTGT | 2520 |
| CAGACAAGAG GTTAATCAGA CAAATAAATG CTACTAGCAA CAAGGAGAAA GACTGCCCTG | 2580 |
| AAAGTGATAA AGACACTATA GAGTAAAGCA AAAATTTTGC AAAACTAATG ACTGTGTATT | 2640 |
| TCAAGACACG ATGATGTTGA AAATCCGCCA AAATCCCAG AAAGATTGT AGAACTTGGG | 2700 |
| GCAGGGTTTC TGAAATCGTG ATGAGTAAAA TCGCCAAAGG GGCAAAAGAT GCATCTGCCA | 2760 |
| CATAATTCAG GAAGGCCAGA TAAAAATCG TATCCCCAAG CGTTGAAATC CACTGGTTGA | 2820 |
| TAGTTAATTG CCTAAATCT CTATTTTGAA GAAATACTTT CATCACAACCT CCTTCTTAAG | 2880 |
| TTCAAATGGG AATCTTTCCC CAAGGATAGA CCGCGATACT ACTAACAACC AAAATTACAG | 2940 |
| TAACATCAAA AGCTGACCAA TGCCATTGTA GACTATATGC AGTCCAATAG GCCAATAAAT | 3000 |
| TGACTTTGTC ATTCTAAATA AGACTGCAAA TATAAGACCT CCACCCATAT AGAAGACAAA | 3060 |
| GTCTGTCAAG ACCCAACCGT GATTACTAAT GTGCGAGACC CCAAATAAAA CAGCGGAACC | 3120 |
| AAGTACATCT AGCCCCATT TCTTTCCTTT TTCCAGAGCA GTCATCACTA ATCCACGATA | 3180 |
| AATCATGTCT TCAAAAATGG GACCTGCAAT CACAGGATAA AAAAAATACA TCAAAAATGC | 3240 |
| TGTAGCCCCCT GTAAAAGTCG GAGCAGCATG TTGATAAGAA ATTTCAATTC GAGTAGGTGG | 3300 |
| GAAAAGAAAA AAGGTAACGA AATTCCAAAC AACAAAAGCA AGCAGAGCTA GGAAGGAATA | 3360 |
| GAAAAGATAG GATCCTTTAA ACTTTCTACT ATTGATTTTC TGCCATTTCC CCGACCAAAT | 3420 |
| CATAGCAATA AGAGCAAATA AAACCACAAG AAAATTCAAC ATCATATCCG ACAGATAATA | 3480 |
| GGCAAAGTCA GATAGCCCAG TAACAAGGTC GCTGCGTAAA ACTAGAACAC TGAACCTCTG | 3540 |
| GTCAGCAATA ACTAGTAGAA AAACATAAT AAAGTAGCGG TGTGAGATTA TCTTTTTCAT | 3600 |
| ATATCACCTT TCTAATATCC AAATACCAAT AAAGTAACAA TGAGTAAGAA ACTATTCCAT | 3660 |
| GAAGCATGCA GAGCTATAGC CCAATAGATG GATCGGGTGT AGCGAAACAT CATACAAAAT | 3720 |
| ATCAAGCCCA TTCCAAAATA CTTTATGAAA TCTGTGTTA TCCAACCATA CTGCAAAACA | 3780 |
| TGCATAGCGC CAAATATGGC AGCGGAAACA AGAACATCAA GATAGTATCT CTTAACCTTA | 3840 |

1028

| | |
|--|------|
| GATAAACTTG TCATCAAAAG ACCACGACAA ACAACCTCTT CTGATACAGG TCGGATAATA | 3900 |
| CTAGTATAAA GTATTTCGCGT AACAAAATAG CTAATTCCTG TTAAATTGGT GGCTACTTCT | 3960 |
| ACGACTGTAC TTCCATTCTG GGTACGAGGA AAGATATAGG TTGTTAGATT TGCCCACACG | 4020 |
| AACAATAAGA AAAAAGAAAG AAGGAAAACA CCCAGGTAAG ACCAACGAAA CTGGAAACGA | 4080 |
| CCCACTCTT TCCAATGTTT ACTTTTGACA AAAGCAATTG TAGCTATAGT TCCCAGAATA | 4140 |
| AGTACCAATA AAACCTGGAA CACATAGTAC ATATTATCAG ACAAGCAAC CATAAAATCT | 4200 |
| AAGCTGATG TGACATTAAA AATGAGGTAA TAAGTCAAAA TCAACAAGCC AGTTGCTAGG | 4260 |
| TGAAATTTCA CTCTTTTCAT TTTCTTCATC CTATTATCTC CTATAAGAGC CTATCTTCTA | 4320 |
| CGGCGGCCAA ACAATCCATC TGCTAAATCT ATAGTCCAAT CAAAAGCTCC ACGATTAGGA | 4380 |
| CTCATCCCTT GATTGCCCCA ACCAGGGTAA ATTCCTGGGA CGCCCCAACC AGATATACCA | 4440 |
| CTTCTCCAC CACCTCCCAT AGAATTTACG AGGTTGCCTC CTCTAACATC TTGCAACTCA | 4500 |
| GCTTCTGTCA ATTCCATTGT TTCTGCAAAT TGTAATTTTA ACATCTTTTA CACTCCTTCA | 4560 |
| ATTATCTTCA TTTGTAAACC ACTTCTGCGA CCTAGGATTT GCTTCAAGTG CTTTACAAGT | 4620 |
| ACAGTATAAC ACGAACATTG GCTTATTTTA GAAATCGCA TATTTGATAT TTTTCTTAT | 4680 |
| AGAAATTTCA GATTTGCGAT TTTGGTGAAT TTGATTACTT CTCTGGTATA ATAAAGTTAC | 4740 |
| TACTAATGAG GAGTGGAGAA ATATGAAGAA ACAAATTTTA ACATTATTGA AA | 4792 |

(2) INFORMATION FOR SEQ ID NO: 157:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2156 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 157:

| | |
|---|-----|
| CCGTTCTCGG CGACGGCCAT CTGATGAAGC TATTTATGAG GGAACTGGC AAGCTGGAGA | 60 |
| GTCAGAGTAT CTAGTCTTTC ACCGATTGCT GTGCAGCAG ATGTGCAGGG AAAAGGAGTT | 120 |
| GCTCAAACCT TCTTAGAGGG CTGATTGAA GGTTTTGATT ATCTTGATTT TCGCTCAGAT | 180 |
| ACGCATGCTG AAAACAAGGT TATGCAACAT ATTTTGTAAA AACTTGGTTT TAAACAAGTC | 240 |
| GGTAAGATGC CAGTAGATGG CGAACGCTTG GCCTATCAAG AATTAAAGAA ATAATGCAAA | 300 |
| AGAAGTATGT AAAATCCTC TACTCCTCAC CAATTGGTAT TCTATCACTT GTAGCTGATG | 360 |
| ACCATTATTT GTATGGAATT TGGGTCAGG AGCAGAAGCA TTTTGAGAGG GGAAGTAGG | 420 |
| ATGAAACGAT AGAAGAAGTT GTTAGTCATC CTATTTTGA CCCAGTTATT GCTTGCTTAG | 480 |

1029

| | |
|--|------|
| ATGATTACTT TAAAGGCAAG CCTCAGGATT TATCCAACTT GCTCTTGGCG CCAATCGGAA | 540 |
| CGAATTTTGA AAAGAGAGTT TGGGACTATT TACAGGGCAT TCCTTATGGT CAGACAGTGA | 600 |
| CCTATGGACA AATTGCTCAA GACCTGCAAG TGGCTTCTGC TCAAGCAATT GGTGGAGCAG | 660 |
| TGGGACGCAA TCCTTGGTCT ATCCTAGTAC CTTGTTCATCG TGTGTTGGGA GCAGGCAAGC | 720 |
| GTCTGACAGG TTATGCTGCA GGAGTGGAAG AGAAAGCTTG GCTCTGGAG CATGAAGGAG | 780 |
| TAGATTTTAA AGATAGAAGC AATAGAAGGA GAAGCACATG TTAGAATTTA TCGAATACCC | 840 |
| CAAATGTTCA ACTTGTAATA AAGCAAAACA AGAATTAAAT CAATTAGGTG TGGACTATAA | 900 |
| AGCCGTCCAT ATCGTGGAAG AAACACCTAG CCAAGAAGTC ATTTTGAATT GGCTAGAAAC | 960 |
| CTCAGGATTT GAATTGAAGC AATTTTTCOA CACCAGTGGT ATCAAATACC GTGAATTAGG | 1020 |
| GCTAAAAGAT AAGGTAGGAA GTTTGTCAAA CCAAGAAGCG GCTGAGTTGC TAGCAAGTGA | 1080 |
| CGGTATGTTG TTAACACGGC CCATTTTAGT AGAAATGGA ACTGTTAAGC AAATCGGTTA | 1140 |
| TCGAAAATCT TATGAGGAAC TGGGACTGAA ATAGTTTTTA TCTATCTCTT TGATAGATAA | 1200 |
| AATATATAAC TTCCCTGTTT CAAAGTATGA TAAACTAGTA GGTAGACAAA GTCTGTATCT | 1260 |
| GACCGTAGCA AATAATTTCA TTGACGGCAG AAGCATGGTA GCATGAATCA TTATCAGAAG | 1320 |
| AGGATGTTTT TATGAATGTT ACAACGATTT TAGCATCAGA TTGGTACCAA AACTTGATGC | 1380 |
| AATTGATFCC GGATGGCAAG CTGTTTAGCC TACGTTCCGT CTTTGATGGA ATCCCTAGAA | 1440 |
| TTGTCCAACA ACTTCCAACA ACAATTATGT TGACAATTGG TGGTGCCTTT TTTGGCTTGG | 1500 |
| TTTTGGCGCT TCTTTTGGCC ATTGTGAAGA TCAATCGTGT CAAGATTTTA TATCCCTTGC | 1560 |
| AGGCCTTCTT TGTTAGTTTC TTAAAAGGGA CACcGATTTT GGTGCAACTC ATGTTGACCT | 1620 |
| ACTACGGAAT CCCTTTGGCT TTGAAAGCCC TCAATCAGCA ATGGGGAACT GGTCTCAATA | 1680 |
| TCAATGCGAT TCCAGCTGCA GCTTTTGCGA TTGTCGCCTT TGCCTTTAAT GAGGCAGCTT | 1740 |
| ATGCTAGTGA AACCATTCGT GCAGCCATTC TCTCAGTTAA TCCTGGTGAG ATTGAGGCGG | 1800 |
| CACGCAGTCT GGGTATGACC CGAGCGCAAG TTTATCGACG AGTGATTATT CCTAATGCAG | 1860 |
| CGGTGGTAGC TACTCCAACC TTGATTAAAT CCCTCATCGG TTTGACCAAG GGAACATCTC | 1920 |
| TAGCTTTTAG TGCGGGTGT GTGGAAGTCT TTGCCCCAAG TCAGATTCTA GGTGGAGCTG | 1980 |
| ATTATCGCTA TTTTGAACGC TTCATCTCCG TTGCCCTTGT TTATTGGGTA GTCAATATCG | 2040 |
| GAATTGAAAG CCTCGGTCGT TTCATCGAGA GAAAAATGGC TATTTCTGCA CCTGATACAG | 2100 |
| TGCAACAGAT GTGAAAGGAG ACCTTCGTTA ATGATTAAGA TTTCGAATTT AAGCAA | 2156 |

(2) INFORMATION FOR SEQ ID NO: 158:

1030

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3140 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 158:

| | |
|--|------|
| GTATCTCTAC ACATGTCTTC AATCGATTTT GTTGTCCTCC AATTTAATTC CTTATATGCT | 60 |
| TTGTCTGCAT TTGCATAACA AGTTGCAACG TCTCCTGAAC GTCTTGGAAC TATTTTATAA | 120 |
| GGAATAGGGA TCTTATTAAC ACTTTCAAAT GTATTTACAA GTTGTAATAC ACTAGTGCCT | 180 |
| TCTCCCGAGC CTAGGTATA GATATAACA TCTGTTTTTT CAGATACTTT TTCTAAAGCT | 240 |
| TTTATATGTC CTATTGCTAA ATCTACTACA TGGATATAAT CACGCACACC AGTACCATCA | 300 |
| AGCGTATCAT AATCATTTC GAACACACTT AGCTCTGATA GCTTACCTAC CGTACTTGT | 360 |
| GCAATATAAG GCATCAAGTT GTTAGGAATT CCTGAGGGAT CTTCCCAAT CAAACCAGAC | 420 |
| TCATGAGCAC CAATTGGATT GAAATAACGA AGCAACGCAA TACTCCATTC TGAATCTGCC | 480 |
| ACATGAACAT CTTTAAAT TTGCTCAAGC ATCACTTTCG TATACCCATA AGGATTGTGTC | 540 |
| GCACCTGTTT GCATCGTCTC AATTAGAGGT GACTGATTGT TAATTCATA TACAGTCGCA | 600 |
| CTTGAAGAAA AGACAATCTT TTTAACATTA AATTCTGACA TCACCTCAAC AAGTGCCAAT | 660 |
| GTACTCATAA TATTATTTTT GTAGTACATC ACAGGCTTTT GCACGGATTC TCCGACAGCT | 720 |
| TTATAACCTG CAAAATGAAT TGCAGCATCA ATCGATTCTT GTTCAAATAC CTTTCTCAAT | 780 |
| GCTTGTTTAT CACAAACATC TAATTCGTAA AACACGGGAC GTATTCTGT AATTGCTTCA | 840 |
| ATACGGTCTA GCACCAAGAT GCTAGAGTTC GAAAGTTGT CGACAATGAT AACTTCTTTT | 900 |
| CCTAAATTTA GTAATTCTAC TACGGTATGG CTACCAATAT AACCAGCTCC GCCTGTTACC | 960 |
| AATATTGCCA TCTGGGTTTC CTCCTAATTA ATTCCAACCG ACTTAACAAA TCTCATAAAC | 1020 |
| GCTTCATGCC CAGACGGTGT ATTCTTATAA ACTCCTGCAT CTTCCAGAAC TCTCGCAAC | 1080 |
| ACTTGTCTG CTTCTGTGTG AACTACGCTA TTAACCTCTT CTTTATTAAT GCGAGGATAT | 1140 |
| TTTTCTTTCA ATTGGTCGGC CCATTCTAAA TGATAATCCG CAATTGCATT ATCCTCTCCT | 1200 |
| AAAAGATATT TTCCAACCTC TTCTAACTCT GGTTCCAAAC GAGGTGGTAA TATCGCAAGT | 1260 |
| CCCATCACTT CGATTAACCC GATATTTTCC TTTTAAATAT GTTGATACATC TTGATGAGGA | 1320 |
| TGGAAACAC CATCTGGGTA TTGTTCACTA GTATGATTAT CTCTTAGAAC AATATCTAAT | 1380 |
| TCGTATCTCC CGTCCACTTT ACGAGCAATA GGAGTACCG TATGGTGTGG GACATCTTCA | 1440 |
| GTCATAGCAA TGATGTCTAC TTCTAAATCT GAATATTCTC TCCACTTATT TAGAATTTTA | 1500 |

1031

| | |
|---|------|
| GTAGCTAAAT CTAACAAGCG ATTTTATTTT TCACTTTGTA ACCTAATTAC TGACATTGGC | 1560 |
| CATTTTACAA TACCAGCATT AACATCCTCA AAGTCTTTAA AACAAAATTC ACTCTCAAAT | 1620 |
| TTTGCTTTTT CCATTGGGAA AATATGTTTC CCTCCCTGGT AGTGGTTATG ACTAAGAATG | 1680 |
| GAGCCTCCTG AGATAGGAAG ATCAGAATTT GAACCAGCAA AATATCCTGG CAAAATATCA | 1740 |
| ACAATCTCCA ATAATTGTTT AAATGTTTTA GAGGTAATAG CCATTGGTAC ATGTTGACTA | 1800 |
| TTCAAAAATA TCGCATGCTC ATTAAAGTAT GAGTAGGGAG AATACTGGAA TCCCCATACT | 1860 |
| TCGTACACAA GTTTCAACCG AATAATTCTA TGATTCGAAC GTGCTGGATA ATTTATTTCG | 1920 |
| CCCTGATATC CTTTATTTC CATACATAGT AAACATTTGG GATAATTAGT TGCTTTTACT | 1980 |
| AATTTTTCAG CAGCAATTGT TTTTGGATCT TTTTCGGGTT TTGACAAATT TATCGTAATC | 2040 |
| TCTAGCTCTC CGTATTTAGT TGATGCTCGA AACTCAATAT TCTTAGCAAT AGCAGAAGTT | 2100 |
| TTAATATAAT CACTATCTTT ACTTAACTTA TAAACTCTT CAACTGCTTC TTGAGGTGAT | 2160 |
| ATATCATATG AACTCCAAA AATATCATTT AATCGACTAG GTAAAGGAAC TATGAAATTC | 2220 |
| ATTAACCTG CTCCTAAACA TTCTTTTCC TCGATTAAAT CTTTAATTTT ACCGTTTTTT | 2280 |
| AAGGCGATTT CCACTAAGTA ATCTTTTATT TGTTTCAGGT CATTTTCATC GGAAATGCGA | 2340 |
| TCAATTCCTT CCTCACCTAT TAACGCTAGT ACTCTATTTT TCACATATAT TTTGTCAATT | 2400 |
| TCATTATACA TTCCGTATTC AATTACTCTA TCAACAAAAT TATCAATAAT TGTTTTCATA | 2460 |
| TATTTTCTT TCTAATTTAT GTTCCCATAT TTTCTATACA TTATCCATTT ATAAATGCT | 2520 |
| TGCGTAGTAT GAGCAATTTT ATCAAGGTGA TGAATAATAT CTAAAGCACT AATTACTTCA | 2580 |
| GAAACGTTCC CATCATCTTC AAATATGTAA TTCATTATTT TCTTTTCCAT ATTTATACTA | 2640 |
| AGCTCTTCTA TCTCATCTG TTTTGTGATA ACAACCATAT CTAAACATCC AGATTGTTCC | 2700 |
| TCTCTATAAC AAGATATAGC CCTATTCATA TGCAGTCCGA TAACTTCATG AAGTATTTTT | 2760 |
| ATTTTGGAAA TAATTTCTT CAAAATTTCA TTATTTTGAA GAATCTGTAG ATTTTAAAA | 2820 |
| ATTTCAACAA TTCTATCCCC AATACGTTCA ATGTCAGTTG ATATTTTAT TACACTAATA | 2880 |
| ATCTTCTTA AGTCATATGA AACAGGATGT TGTAACAAA TTAACCTATA TCCTTTTTTA | 2940 |
| TCAATATTTA GAACTGACTC ATTTATGATT AAATCTTCTT TAATCAATTC TACTCGTTCT | 3000 |
| TCATTTGATA AATATTCAAA TAACTTCTCA TATTTATCAA GCACAGATAC CCAAATGGTC | 3060 |
| TCTAAATTAT TTGATAATTC TATAATTTCA TTTTCTAAAT ATAACCTTAA CATTTAGGTA | 3120 |
| CCTCTTCTTA ACAAAGTTCTG | 3140 |

(2) INFORMATION FOR SEQ ID NO: 159:

1032

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 9048 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 159:

| | |
|---|------|
| CCGGATGATT TCCTGGTCAG ATAGGGGGAA AGTGACTTCC TCAGCAATCG CGCGTAGAGT | 60 |
| AGGATTCCCT TCACGGATAA TATCGTTCAT ATCAATTAAG TGAGCAGCTT TTGTAATACG | 120 |
| TTCTATTGCA GACATTTTCT CTCCTTATAT TATGTTTAGT GCAGTTAGCT ACTGCCAAAG | 180 |
| CCCAAGTGGT ATACTTGGA TAAGCCACTG TGGATTAGTT CATTTTCTTT CATTACCTCT | 240 |
| ACATGATATC ACAAATGAC AAGAATTGAA AGCATTATGG CATTTAGGAT TTATAGAAAA | 300 |
| TAGATAGGAA GTTCAATTCA ATTGTGAAAG AAATACTTAT CTGTGATATA ATAAAAAGAA | 360 |
| AAGGCTTGCA TAAGAAAGTA GGGAGAACGA AGATACAAAG AAGACAAAAT CGAAATCAGG | 420 |
| GTGGTTTAGC TTTTCGTTTT ATGAAGGGCT TGGTAACTT TTTAGGAGTT ATCGCAAGTG | 480 |
| GAGCAATAAG GGATTGTGG CGATACTCTT GCTAGCAGTT GGTTTATCAA TGGGCTTGTT | 540 |
| CTTGTTGTTT GAAAGCTTCC AAGGAATCCC TTGACTAGTC AAAACGAGA TACTATTTCT | 600 |
| CAAGAGGGGA CTAAGCAAAA GTCTCAGGAG TAGGAAGAGG AAAAACTGC CAGAATTATG | 660 |
| GCCACGGGG ATTTGCTCTA CCACGATGGA CTTTCTTTT CAGCTAAAAA AGAAGACGGT | 720 |
| ACCTATGACT TTCATGAAAA TTTTGAGTAT GTGACTCCTT GGCTCAAGCA AGGGGACTAA | 780 |
| GCAGCAGATT TAGCTATTGG TGATTTTGAA GGAACCATTA ATAAGGATCA TTATTTAGCG | 840 |
| GGTTATCTTC TCTTAAATGC TCCTGTGAA GTTATGGATG CTATTAAGGA GGCAGGTTAT | 900 |
| CATGTGCTGG ATTTAGCTCA TAATCATATT TTGGATTCGC AAATTGAGGG AGTTATTTCA | 960 |
| ACGGCCGATA TTATTGAGAA AGCTGGAATC ACTCCAATCG GAGTTTATAC GCACGAACCA | 1020 |
| CGTGATCAGG CTCGCTGGT CATTAAAGGAA GTGAATGGTA TCAAGGTTGC ATTGTTAGCC | 1080 |
| TATTCCTATG GTTCAATGG AATTGAGCAG TATATTTCTC AGGAAGACTA TAATCGTTAT | 1140 |
| CTTTCAGATT TAAACGAAGA TAAGATGAAG GTTGAAATTG AACGGGCAGA GAAGGAAGCA | 1200 |
| GATATCACCA TTATCATGCT TCAGATGGGT GTTGAGTATC GATTGGAACC AACTGAAGAA | 1260 |
| CAAAAAGCTC TTTATCACAA GATGATCGAT TTGGGAGCGG ATATTATCTT TGGAGGGCAT | 1320 |
| CCTCACGTTG TTGAACCATC TGAAACGGTT GAAAAAGATG GAGATAAGAA ACTCATTATC | 1380 |
| TATTAAATGG GGAACCTCAT TTCCAATCAA CGAATTGAAT CTATGGGAGA TGAAGAGAAT | 1440 |
| GCTAAGTGA CTGAACGTGG TGTCTCATG GATGTCACCA TCAAGAAGAA GGATGGAAAA | 1500 |

1033

| | |
|--|------|
| ACAACTATCG GAACAGCTAA AGCTCATCCT ACTTGGGTCA ATCGAACACC AAAGGGAACC | 1560 |
| TTTTCACCAG AAGGATATCC CTTGTATCAT TACCAAACCT ATATTTTGA AGATTTTATA | 1620 |
| GAGGATGGCA GTCATCGTGA CCAGTTAGAT GAAGCGACTA AGGAACGAAT TGATACAGCC | 1680 |
| TATAAAGAAA TGAATGAACA TGTGGGATTG AAGTGGTATT AGCTTGAATC CAGAGGAAAG | 1740 |
| TAAATGATGA TTAAGGTAAT TGCGACAGAT ATGGATGGGA CCTTGCTGGA TGCTAGAGGT | 1800 |
| CAGCTTGATC TCCCACGATT GGAAAAGATT TTAGATCAGT TGGATCAAAG GGGCATTTCGT | 1860 |
| TTTGTCAATG CGACGGGCAA TGAAATTCAC CGCATGAGAC AACTACTGAG TCCCTTGGTG | 1920 |
| GATCGAGTGG TTCTGGTTGT TGCTAATGGC GCTCGTATTT TTGAAAACAA TGAATTGATT | 1980 |
| CAGGCTCAGA CATGGGATGA CGCCATTGTC AACAAAGGCTT TGACTCATTT CAAGGTCGA | 2040 |
| GCGTGTCAAG ACCAGTTTGT TGTAACGGGG ATGAAGGGTG ATTTTGTCAA GGAAGGTACG | 2100 |
| ATTTTACAG ATCTTGAAAG TTTTATGACT CCAGAAATGA TTGAAAAATT CTACCAACGG | 2160 |
| ATGCAATTG TGGATGAATT AACATCTGAC CTCTTTGGTG GTGTGCTCAA GATGAGCATG | 2220 |
| GTTGTTGGTG AGGAACGTTT GAGTTCGGTT TTGGAAGAAA TCAATGCTCT CTTTGATGGC | 2280 |
| CGTGTCGAG CTGTATCCAG TGGCTATGGT TGCATGATA TCCTCCAAGC TGGGATTCAT | 2340 |
| AAAGCATGGG GCTTGAGGA ATTACTCAAG CGCTGGGACT TGAAATCCCA AGAAATCATG | 2400 |
| GCTTTTGGTG ATAGTGAAAA TGATGTTGAA ATGCTTGAAA TGGCTGGAAT TGCCTATGCG | 2460 |
| ATGGAAAATG CTGATGAGAA AGCCAAAGCT GTGGCGACTG CTCTAGCACC AGCCAAACAGC | 2520 |
| CAAGGAGGAG TTTATCAAGT CTTGGAAC TGGTTAGAAA AAGGAGAATG AAGTGGCAGT | 2580 |
| ACAGTTATTA GAAAATTGGC TCCTAAAGGA ACAAGAAAA ATTCAAACTA AGTATCGTCA | 2640 |
| CCTAAATCAC ATTTCTGTTG TAGAACCAAA CATTCTTTTT ATGGGGATT CCATTGTCGA | 2700 |
| GTATTATCCT CTACAGGAGC TATTTGGGAC TTCAAAGACG ATGTGCAATC GAGGAATTCTG | 2760 |
| TGGCTATCAG ACAGGACTGT TACTAGAGAA CCTTGATGCT CATCTATATG GTGGAGCAGT | 2820 |
| AGATAAAATT TTTCTTCTGA TTGGGACAAA TGATATCGGA AAGGATGTTT CTGTGAATGA | 2880 |
| GGCTCTCAAT AATCTCGAAG CTATCATCA ATCCGTTGCT CGCGATTATC CATTGACAGA | 2940 |
| GATTAAATTG CTTTCCATTT TGCTGTCAA TGAGAGAGAG GAGTACCAGC AGGCAGTCTA | 3000 |
| TATCCGCTCG AATGAAAAA TTCAGAACTG GAATCAAGCC TATCAAGAGC TTGCATCTGC | 3060 |
| CTATATGCAG GTGGAATTTG TGCCAGTATT TGATTGTTTG ACAGACCAAG CAGGCCAACT | 3120 |
| CAAAAAGAA TATACAACTG ATGGACTGCA CCTCAGTATT GCTGGTTATC AGGCTTTGTC | 3180 |
| AAAATCCTTG AAAGACTATC TTTACTAAAT AGCTAAATAA TGTAAATTT GAGCATAATA | 3240 |

1034

| | |
|--|------|
| TCTTGTAATAA AATTCTAAAA TCCTTTAAAA TAAAAAGTGA CGGAGGAATT TATGAATGTA | 3300 |
| AATCAGATTG TACGGATTAT TCCTACTTTA AAAGCTAATA ATAGAAAATT AAATGAAACA | 3360 |
| TTTATATTG AAACCCTTGG AATGAAGGCC TTGTTAGAAG AATCGGCCTT TCTGTCACTA | 3420 |
| GGTGACCAAA CGGGTCTTGA AAAGCTGGTT TTAGAAGAAG CTCCCAGTAT GCGTACTCGT | 3480 |
| AAGGTAGAGG GAAGAAAAAA ACTAGCTAGA TTGATTGTCA AGGTGGAAAA TCCCTTAGAA | 3540 |
| ATTGAAGGAA TCTTATCTAA AACAGATTCG ATTCATCGAT TATATAAAGG TCAAAATGGC | 3600 |
| TACGCTTTTG AAATTTTCTC ACCAGAAGAT GATTGTGATT TGATTTCATGC GGAAGATGAC | 3660 |
| ATAGCAAGTC TAGTAGAAGT AGGAGAAAAG CCTGAATTTC AAACAGATTT GGCATCAATT | 3720 |
| TCTTTAAGTA AATTGTGAGT TTCTATGGAA TTACATCTCC CAACTGATAT CGAAAGTTTC | 3780 |
| TTGGAATCAT CTGAAATTGG GGCATCCCTT GATTTTATTC CAGCTCAGGG GCAGGATTTG | 3840 |
| ACTGTGGACA ATACGGTTAC CTGGGACTTA TCTATGCTCA AGTTCCTGGT CAATGAATTA | 3900 |
| GACATAGCAA GTCTTCGCCA GAAGTTTGAG TCTACTGAAT ATTTTATTCC TAAGTCTGAA | 3960 |
| AAATCTCTCC TTGGTAAAGA TAGAAATAAT GTTGAATTGT GGTGTGAAGA AGTATGAAGT | 4020 |
| GGACCAAGAT TATTAATAAA ATAGAAGAAC AAATCGAGGC AGGGATTAT CCCGGAGCCT | 4080 |
| CTTTTGCGTA TTTTAAGGAC AATCAATGGA CAGAGTTCTA TTTAGGCCAG AGTGACCCAG | 4140 |
| AGCATGGCTT GCAGACTGAG GCAGGACTAG TTTATGACCT AGCTAGTGTC AGCAAGGTTG | 4200 |
| TTGGGGTTGG CACAGTTTGT ACCTTCTTGT GGGAAATAGG TCAATTAGAT ATTGATAGAC | 4260 |
| TGGTAATAGA TTTTTTACCT GAGAGTGATT ATCCAGACAT CACTATTTCG CAGCTCTTGA | 4320 |
| CTCATGCAAC AGACCTTGAT CCTTTTATTC CTAATCGTGA TCTTTTAACA GCCCCTGAAT | 4380 |
| TAAAGGAAGC GATGTTTCAT CTCAACAGAC GAAGTCAGCC AGCCTTTCTT TATTCGGATG | 4440 |
| TCCATTTTTT GCTGTTGGGC TTTATTTTGG AAAGAATTTT TAATCAAGAT TTGGATGTGA | 4500 |
| TTTTAAAGGA TCAAGTCTGG AAACCTTGGG GAATGACGGA AACTAAGTTT GGGCCAGTTG | 4560 |
| AGCTTGCTGT TCCAACAGTT AGAGGTGTAG AGGCAGGCAT AGTGCATGAT CCCAAGGCTC | 4620 |
| GTCTCCTGGG TAGACATGCT GGGAGTGCTG GTTTATTTTC GACTATAAAG GATTTACAAA | 4680 |
| TCTTTTATGA AACTATTTA GCAGATGATT TTGCAAGAGA CTTAAATCAA AATTTTCTC | 4740 |
| CTTTGGATGA CAAGGAACGT TCTTTAGCAT GGAATTTGGA AGGAGATTGG CTAGACCATA | 4800 |
| CGGGCTATAC AGGTACCTTT ATCATGTGGA ATCGTCAGAA GCAAGAAGCC ACTATTTTCC | 4860 |
| TATCGAATCG TACCTATGAA AAGGACGAGA GAGCTCAATG GATATTAGAC CGCAATCAAG | 4920 |
| TGATGAACTT GATTCGCAAA GAAGAGTAAG GAGAGACATG TCAAATAGTT TAAAGGGAC | 4980 |
| TTTACTAACA GTTGTGGCTG GTATTGCTTG GGGGTTGTCA GGAACGAGTG GCCAATACCT | 5040 |

1035

| | |
|--|------|
| AATGGCACAC GGAATTTTCGG CTCTGGTCTT GACTAACTTG CGTCTTTTAA TCGCTGGTGG | 5100 |
| AATTCTCATG CTCTTGGCTT ATGCTACTGC AAAGGATAAA ATACTGGTCT TTTTAAAGGA | 5160 |
| TAGAAAGAGT TTGCTGTCTC TTCTTATTTT TGCTCTGATT GGTCTTTTTC TCAACCAATT | 5220 |
| CGCCTATCTG TCTGCTATTC AGGAGACCAA TCGGGGAACA GCGACGGTGC TTCAGTATGT | 5280 |
| TTGTCCTGTC GGAATTTTAA TTTATAGCTG TATCAAGGAT AGGGTGGCAC CGACACTGGG | 5340 |
| AGAGATAGTT TCCATCATAT TCGCCATCGG AGGAACCTTC CTGATCGCAA CACATGGGCA | 5400 |
| GTTGGACCAG TTATCCATGA CACCTGCTGG TCTGTTCTGG GGTCTCTTTT CTGCCTTGAC | 5460 |
| TTATGCTCTG TATATCATTT TACCCATAGC CTGATTAAA AAGTGGGGGA GCAGCTTGGT | 5520 |
| CATTGGTGTG GGAATGGTCA TAGCAGGTTT GGTGCGCCCTT CCTTTTACAG GGGTTCTACA | 5580 |
| GGCCGATATC CCGACTAGTC TTGATTTTCT CCTTGCCTTT GCAGGCATTA TCCTTATCGG | 5640 |
| GACTGTCTTT GCCTATACAG CTTTCCTTAA AGGAGCCAGT CTGATAGGAC CGGTCAAGTC | 5700 |
| AAGCTTGTG GCTTCAATTG AGCCAATATC GCGGATTTTC TTTGCCCTCT TAATAATGAA | 5760 |
| TGAACAATTT TATCCCATG ATTTTCTTGG TATGGCAATG ATATGTGTTG CTGTAACTTT | 5820 |
| GATTTCTTTG AAAGATTTAT TCTTAGAAAA ATAAAAAGA CTCTTTGTCC GTGACAGAGA | 5880 |
| GTTTTTGCGT GGTAATCTAA TTATTTTCAA GATAAAATTC AAAGCGTTCG CCTACATATT | 5940 |
| GACTTTTAC GTATTCAAAA GCAGTACCAT CTTCTAGGTA GGAAACCTGG GTCAATCCAA | 6000 |
| GAATAGCATG TCCTTTTCA ACTTCCAAAT AGTGGGCAAT CTTTCTTTTA GCAAGGCGAG | 6060 |
| CATAGATGGT CTGTTGAGAT TTGCCGATAC GATAGCCATG TTTTGCAGG GTTTGGAAGA | 6120 |
| AATGACTGGT GATTTCTTCT TTTTAAAGT CCTTAATGAA TTTTTCAGGA ATAGAAGCAA | 6180 |
| CTTCATAAAC TAGGGGAAC TGGTCGGCAT AGCGGACCCG CTCCATTCGG ATAATATTGT | 6240 |
| CCGTTGGAAA AATTCCTAGC TTGGCAACTT CTTGCTCATT GGAATGGTT TTTTGTAGG | 6300 |
| AAATGAGCTG GCTAGAGGGA ACTTTACCTT GGGATTTGAC AATTTTCAGTA AAAGTGGTTG | 6360 |
| TCCCTCGCAT CTTTCTTGT ACTCGAGTAC TGGAACAAA GGTGCCGCTT CCTACACGGC | 6420 |
| GCTCTAAGAC GCCTTCTTCG ACTAATAGAG ATACGGCTTG GCGGAGGGTC ATGCGACTGA | 6480 |
| CCGCAAACTG CTCAGCTAAA TCTCTTTCAC TGGGAAGCCT CTCACCAATA GCCCAACGGT | 6540 |
| ACTCGTCAAT ATCCTTTTTT ATCTGATCAT GGATTTTTAT ATAAGCAGGT AGCATATTTT | 6600 |
| TCACTTCATT TCTATCTTTT CTCTATTGTA CCCCAATAAA CTAGAAAAAG TCAAACTTCG | 6660 |
| CCTTGTTTAG TTGGTAATTC GCCCTTATTT GTGATAGAAT ATTGAGAAAA GATATTTCTT | 6720 |
| TTGAGAAAGG AAAAAGATGA GCAACATTC AACTGATTTG CAAGATGTAG AAAAAATCAT | 6780 |

1036

| | |
|--|------|
| CGTATTGGAC TATGGTAGCC AGTACAACCA GCTGATTTCA CGCCGTATCC GTGAGATTGG | 6840 |
| TGTTTTTTTCA GAACTAAAAA GCCATAAAAT TTCAGCTGCT GAAGTTCGTG AAGTCAATCC | 6900 |
| TGTAGGAATT ATTCTATCAG GTGGTCCAAA TTCTGTATAT GAAGATGGTT CATTTGATAT | 6960 |
| TGACCCAGAA ATCTTCGAAC TCGGAATTCC AATTTTGGGA ATCTGTTATG GTATGCAGTT | 7020 |
| ATTGACCCAT AAAC TTGGAG GAAAAGTTGT TCCTGCAGGT GATGCTGGAA ATCGTGAATA | 7080 |
| CGGTCAATCA ACCCTAACTC ACACACCATC AGCGCTTTT GAATCAACAC CTGATGAACA | 7140 |
| GACTGTTTTG ATGAGCCATG GTGATGCGGT TACTGAGATT CCTGCTGACT TTGTTCGTAC | 7200 |
| AGGTACATCA GCTGACTGCC CATACGCAGC CATCGAAAAC CCAGATAAAC ACATTACGG | 7260 |
| TATCCAATTC CACCCAGAAG TTCGTCATTC TGTATACGGA AATGATATCC TTCGTAACCT | 7320 |
| TGCCCTTAAC ATTTGTAAGG CTAAAGGTGA CTGGTCAATG GATAATTTCA TTGACATGCA | 7380 |
| GATCAAAAAA ATTCGTGAAA CCGTCGGTGA TAAACGTGTC CTTCTTGGTC TATCAGGTGG | 7440 |
| TGTTGACTCA TCTGTCGTTG GGGTTCTTCT CCAAAAAGCG ATTGGCGATC AATTGATCTG | 7500 |
| TATCTTCGTA GACCACGGTC TTCTTCGTAA AGGCGAAGCT GATCAAGTTA TGGACATGCT | 7560 |
| CGGTGTAAG TTTGGTTTGA ATATCGTCAA AGCAGACGCT GCTAAACGTT TCCTTGACAA | 7620 |
| ACTTGCTGGC GTTCTTGACC CTGAACAAAA ACGTAAATC ATCGGTAACG AGTTTGCTA | 7680 |
| TGTATTCGAT GACGAAGCAA GCAAGCTCAA AGATGTGAAA TTCCTTGCTC AAGTACTTT | 7740 |
| ATATACAGAT GTTATCGAGT CTGGTACGGA TACAGCTCAA ACTATCAAGT CACACCACAA | 7800 |
| CGTGGGGTTC TTCCAGAAGA TATGCAGTTT GAATTGATTG AACCCTCAA TACTCTTTAC | 7860 |
| AAGGATGAAG TTCGTGCTCT TGGTACAGAG CTTGGTATGC CAGACCATAT CGTATGGCGC | 7920 |
| CAACCATTCC CAGGACCAGG ACTTGCTATC CGTGTATGG GTGAAATCAC TGAAGAGAAA | 7980 |
| CTTGAAACCG TTCGTGAATC AGACGCTATT CTTGCTGAAG AAATCGCTAA AGCTGGACTT | 8040 |
| GACCGCGATA TTTGGCAATA CTTCACTGTT AACACAGGCG TTCGTTCACT CGGTGTTATG | 8100 |
| GGTGACGGTC GTACGTATGA CTACACGATT GCAATCCGTG CTATCACTTC TATCGATGGT | 8160 |
| ATGACTGCTG ATTTTGCCAA AATTCCATGG GAAGTACTTC AAAAAATCTC AGTACGTATC | 8220 |
| GTAAATGAAG TGGATCATGT TAACCGTATC GTCTACGATA TTACAAGTAA ACCACCTGCA | 8280 |
| ACAGTTGAGT GGAATAATC GCAAAAAAAT TAAAAGCTTT GTAAATCAA CGGTTACAGA | 8340 |
| GGATTAAAAA CTGTAACCTGG GATTAAAACG GGAACATTTG CTAAAAAGAA TAAATTGAAT | 8400 |
| AATAGTTCCA AGTGGTTTAC ATTTGGACAA AAAATTAGAC CGTAGTTTTC AAGCTGCGGT | 8460 |
| CTTTTGATAT ATATAATGAG AATTAATGGC TCTTTGTCAA CTGTAGTGGG TTGAAGTCAG | 8520 |
| CTAAGCTCGA GAAAGGACAA ATTTTGTCTT TTCTTTTTTG ATATTCAGAG CGATAAAAAAT | 8580 |

1037

| | |
|---|------|
| CCGTTTTTTG AAGTTTTCAA AGTTCCGAAA ACCAAAGGCA TTGCGCTTGA TAAGTTTGAT | 8640 |
| GAGATTATTG GTCGCTTCCA ATTTGGCGTT AGAATAGTGT AGTTGAAGGG CGTTGACGAT | 8700 |
| TTTCTCTTTG TCCTTTAGAA AGGTTTTAAA GACAGTCTGA AAAAGAGGAT GAACCTGCTT | 8760 |
| TAGATTGTCC TCAATGAGTC CGAAAAATTT CTCCGCTTCC TTATTCTGAA AGTGAAACAG | 8820 |
| CAAGAGTTGA TAGAGCTGAT AGTGATGTTT CAAGTCTTGT GAATAGCTCA AAAGCTTGTT | 8880 |
| TAAAATCTCT TTATTGGTTA AATGCATACG AAAAGTAGGG CGATAAAAAT GTTTATCGCT | 8940 |
| GAGTTTACGA CTATCCTGTT GTATGAGCTT CCAGTAGCGC TTGATAGCCT TGTATTCATG | 9000 |
| AGACTTTCGA TCCAATTGAT TCATGATTTC AACACGCACA CGACTCGG | 9048 |

(2) INFORMATION FOR SEQ ID NO: 160:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10399 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 160:

| | |
|---|-----|
| GTACCTTTAT TGATGAATGG ACTGTTTAAA TCAGTAGCAC GCCAACCAGA TATGCTTTCT | 60 |
| GAGTTTCGTA GTTTGATGTT TTTAGGTGTT GCCTTTATTG AAGGAACTTT CTTTGTAAC | 120 |
| CTTGCTCTTCT CATTATTAT CAAATAAATA CATGGAACGA GAAGAAAAGG GAGGATTTTA | 180 |
| GATGGAAGAA AGTATTAATC CAATCATCTC TATTGGTCCT GTTATCTTCA ATCTGACTAT | 240 |
| GTTAGCCATG ACTTTGTTGA TTGTGGGAGT TATTTTGTG TTTATTTATT GGGCAAGCCG | 300 |
| CAATATGACC TTGAAACCCA AAGGAAAGCA AAATGTACTT GAGTATGTCT ATGACTTTGT | 360 |
| TATTGGATTT ACAGAACCTA ACATTGGTTC GCGCTACATG AAAGATTACT CACTCTTTT | 420 |
| CCTTTGTTTA TTCCTTTTCA TGGTGATTGC CAATAACCTT GGCTTAATGA CAAAGCTTCA | 480 |
| AACGATCGAT GGGACTAACT GGTGGAGTTC GCCAACCGCT AATTTACAGT ATGACTTAAC | 540 |
| CTTATCTTTT CTTGTCATTT TGTGACACA TATAGAAAGC GTTCGTCGTC GTGGATTTAA | 600 |
| AAAAAGTATA AAATCTTTTA TGAGTCTGT TTTTGTGATA CCGATGAATA TCTTGGAAGA | 660 |
| ATTTACAAAC TTCTTATCTT TGGCTTTGCG GATTTTGGG AATATCTTTG CAGGAGAGGT | 720 |
| CATGACGAGT TTGTTACTTC TTCTTTCCCA CCAAGCTATT TATTGGTATC CAGTAGCCTT | 780 |
| TGGAGCTAAT TTGGCTTGGA CTGCATTTTC TGTCTTTATT TCCTGCATCC AAGCTTATGT | 840 |
| TTTACTCTT TTGACATCTG TGTATTTAGG GAATAAGATT AATATTGAAG AGGAATAGAA | 900 |

1038

| | |
|--|------|
| AGGAGTAACT GATGCACGTA ACAGTAGGTG AATTAATTGG TAATTTTATT TTAATCACTG | 960 |
| GCTCTTTTAT TCTTTTGCTA GTCTTGATTA AAAAATTTGC ATGGTCTAAT ATTACAGGCA | 1020 |
| TTTTCGAAGA AAGAGCTGAA AAAATTGCTT CAGATATTGA CAGAGCTGAA GAAGCCCGTC | 1080 |
| AAAAAGCAGA AGTATTGGCT CAAAAACGCG AAGATGAATT GGCTGGTAGC CGTAAAGAAG | 1140 |
| CTAAGACAAT CATTGAAAAT GCAAAGGAAA CAGCTGAGCA AAGTAAGGCT AATATCTTAG | 1200 |
| CAGATGCTAA ACTAGAAGCA GGACACTTAA AAGAAAAAGC CAATCAAGAA ATTGCTCAAA | 1260 |
| ATAAAGTAGA AGCTTTACAG AGTGTTAAGG GTGAGGTCGC AGATTTGACC ATCAGCTTAG | 1320 |
| CTGGTAAAAAT CATCTCACA AACCCTTGACA GTCATGCCCA TAAAGCACTC ATTGATCAGT | 1380 |
| ATATCGATCA GCTAGGAGAA GCTTAATGGA CAAGAAAACA GTAAAGGTAA TTGAAAATA | 1440 |
| CAGCATGCCT TTTGTCCAAT TGGTACTTGA AAAAGGAGAA GAAGACCGTA TCTTTTCAGA | 1500 |
| CTTGACTCAA ATCAAGCAAG TTGTTGAAAA AACAGGTCTG CCTTCTTTT TAAAAAAGT | 1560 |
| GGCAGTAGAC GAGTCGGATA AGGAAAAAAC AATTGCTTTT TTCCAAGATT CTGTGTCGCC | 1620 |
| TTTATTACAA AACTTTATCC AGGTCTTGGC CTACAATCAC AGAGCAAATC TTTTTTATGA | 1680 |
| TGTGCTTGTA GATTGCTTGA ACCGACTTGA AAAAGAAACA AATCGATTTG AAGTGACGAT | 1740 |
| TACGTCTGCT CATCCTCTAA CTGATGAACA GAAGACTCGT TTGCTCCCTT TGATTGAGAA | 1800 |
| AAAAATGTCT CTGAAAGTAA GGAGTGTAAG AGAACAAATC GATGAAAGTC TCATTGGTGG | 1860 |
| TTTTGTCAAT TTTGCCAATC ACAAGACAAT TGATGTGAGT ATTAAACAAC AACTTAAAGT | 1920 |
| TGTTAAAGAA AATTTGAAAT AGAAAGTGGT GTTCTTTTGG CAATTAACGC ACAAGAAATC | 1980 |
| AGCGCTTTAA TTAAGCAACA AATTGAAAAT TTCAAACCCA ATTTTGATGT GACTGAAACA | 2040 |
| GGTGTTGTAA CCTATATCGG GGACGGTATC GCGCGTGCTC ACGGCCTTGA AAATGTCATG | 2100 |
| AGTGAGAGT TGTGAATTT TGAAAACGGC TCTTATGGTA TGGCTCAAAA CTTGGAGTCA | 2160 |
| ACAGACGTTG GTATTATCAT CCTAGGTGAC TTTACAGATA TCCGTGAAGG CGATACAATC | 2220 |
| CGCCGTACAG GGAAAATCAT GGAAGTCCCT GTAGGTGAAA GTCTGATTGG TCGTGTGTG | 2280 |
| GATCCGCTTG GTCGTCCAGT TGACGGTCTT GGAGAAATCC AACTGATAA AACTCGTCCA | 2340 |
| GTAGAAGCAC CAGCTCCTGG TGTATGCAA CGTAAGTCTG TTTCAGAACC ATTGCAAAC | 2400 |
| GGTTTGAAAG CTATTGACGC CCTTGTACCG ATTGGTCGTG GTCAACGTGA GTTGATTATC | 2460 |
| GGTGACCGTC AGACAGGGAA AACAACCATT GCGATTGATA CAATCTTGAA CCAAAAAGAT | 2520 |
| CAAGATATGA TCTGTATCTA CGTCGCGATT GGACAAAAAG AATCAACAGT TCGTACGCAA | 2580 |
| GTAGAAACAC TTCGTACGTA CGGTGCCTTG GACTACACAA TCGTTGTGAC AGCCTCTGCT | 2640 |
| TCACAACCAT CTCCATTGCT CTTCTAGCT CTTATGCTG GGGTTGCTAT GCGGAAGAA | 2700 |

1039

| | |
|---|------|
| TTTATGTATC AAGGTAAGCA TGTTTTGATT GTATACGATG ATCTATCAAA ACAAGCGGTA | 2760 |
| GCTTATCGTG AACTGTCGCT CTTGCTTCGT CGTCCTCCAG GTCGTGAAGC CTTCCCAGGG | 2820 |
| GATGTTTTCT ATCTCCACAG CCGTTTGCTT GAGCGCTCAG CTAAAGTTTC TGATGAACTT | 2880 |
| GGTGGTGGAT CAATTACAGC CCTACCATTT ATCGAGACAC AAGCAGGAGA TATCTCAGCC | 2940 |
| TATATCGCAA CCAACGTGAT TTCTATCACT GATGGACAAA TCTTCCTTGG CGATGGCCTC | 3000 |
| TTCAATGCAG GTATTCTGCC AGCCATCGAT GCGGGTTCAT CTGTATCTCG TGTAGGTGGT | 3060 |
| TCTGCACAAA TCAAAGCCAT GAAGAAGGT GCTGGTACAC TTCGTATCGA CCTTGCCTCA | 3120 |
| TACCGTGAGT TGGGAAGCCTT TACTAAGTTT GGTCTGACT TGGACGCAGC AACACAGGCT | 3180 |
| AAGTTGAACC GTGGACGTCG TACCGTTGAG GTCTTGAAAC AACCTGTTCA CAAACCATTA | 3240 |
| CCTGTTGAGA AACAAGTAAC CATTCCTTAT GCTTTGACAC ATGGTTTCTT GGATACTGTT | 3300 |
| CCAGTAGATG ATATTGTTTCG TTTCGAGGAA GAGTCCATG CCTTCTTTGA TGCTCAACAT | 3360 |
| CCAGAGATTT TGGAAACCAT TCGTGATACA AAAGACTTGC CAGAAGAAGC AGTCTTGGAT | 3420 |
| GCTGCGATTA CAGAGTTTCT CAATCAATCT AGCTTCCAAT AAGAATAGAG GTGTCAGATG | 3480 |
| GCAGTATCTC TAAATGATAT TAAAACAAAA ATCGCCTCAA CAAAAATAC GAGTCAAATC | 3540 |
| ACTAATGCCA TGCAAATGGT ATCGGCTGCT AAGCTAGGTC GTTCTGAAGA AGCTGCTCGC | 3600 |
| AACTTCCAAG TTACGCTCA GAAAGTCCGT AAAGTTTGA CAGATATCCT TCATGGTAAT | 3660 |
| GGAGCTGGTG CTTCAACTAA TCCGATGTTG ATTAGCCGTT CTGTGAAGAA GACAGGCTAT | 3720 |
| ATCGTTATCA CTTCAGACCG CGGTTGGTT GGAGGTTATA ATTCCTCTAT TTTGAAAGCT | 3780 |
| GTTATGGAGT TGAAAGAAGA ATACCACCCA GACGGTAAAG GTTTTGAAAT GATCTGTATC | 3840 |
| GGTGGGATGG GAGCTGATTT CTTTAAGGCT CGCGGTATTC AACCACTTTA TGAATTACGT | 3900 |
| GGCTTGTCAG ACCAACCTAG CTTTGATCAA GTTCGTAAGA TTATTTCAAA AACTGTTGAA | 3960 |
| ATGTACCAAA ATGAACCTT TGATGAGCTT TATGTTTGCT ACAACCACCA TGTCATACG | 4020 |
| CTAACCAGTC AAATGCGTGT GGAACAAATG CTTCCGATTG TTGACTTGGA TCCAAATGAA | 4080 |
| GCGGATGAAG AGTACAGCTT GACTTTTGAA TTGGAAACCA GCCGAGAAGA AATTCTGGAG | 4140 |
| CAGTTGTTGC CTCAGTTTGC AGAAAGTATG ATTTACGGTG CCATTATCGA TGCCAAGACA | 4200 |
| GCTGAGAATG CTGCGGCAT GACAGCCATG CAAACAGCGA CAGATAATGC TAAGAAAGTC | 4260 |
| ATCAATGATT TGACAATTCA GTATAACCGT GCCAGACAGG CGGCGATTAC ACAAGAAATT | 4320 |
| ACAGAAATCG TAGCAGGTGC TAGTGCCTTA GAATAGGCTC TAGTCCAGCT CGTATGAAAA | 4380 |
| TGAACCTAGG ACCTAGTTGA GCTAGGAACC GACAGTATCT TATATAGAAT AGGAGAAGGA | 4440 |

1040

| | |
|--|------|
| GATGAGTTCA GGTAAAATTG CTCAGGTTAT CCGTCCCGTT GTAGACGTTT TGTTCGACG | 4500 |
| AGGGGAAAAA CTCCTGAGA TTAACAATGC ACTTGTCGTC TACAAAAATG ACGAAAGAAA | 4560 |
| AACAAAAATC GTCCTGAAG TAGCCTTGA GTTAGGAGAT GGTATGGTTC GTACTATCGC | 4620 |
| CATGGAATCA ACAGATGGGT TGACTCGTGG AATGGAAGTA TTGGACACAG GTCGTCCAAT | 4680 |
| CTCTGTACCA GTAGGTAAAG AAACTTTGGG ACGTGTCTTC AACGTTTGG GAGATACCAT | 4740 |
| TGACTTGGAA GCTCCTTTTA CAGAAGACGC AGAGCGTCAG CCAATTCATA AAAAAGCTCC | 4800 |
| AACTTTGTAT GAGTTGTCTA CCTCTTCTGA AATCCTTGAA ACAGGGATCA AGGTTATTGA | 4860 |
| CCTTCTTGCC CCTTACCTTA AAGGTGGTAA AGTTGGACTT TTCGGTGGTG CCGGAGTTGG | 4920 |
| TAAAACTGTC TTAATCCAAG AATTGATTCA CAACATTGCC CAAGAGCAGC GTGGTATTTT | 4980 |
| AGTATTTGCT GGTGTTGGGG AACGTACTCG TGAGGGGAAT GACCTTTACT GGGAAATGAA | 5040 |
| AGAATCAGGC GTTATCGAGA AAACAGCCAT GGTCTTTGGT CAGATGAATG AGCCACCAGG | 5100 |
| AGCACGTATG CGTGTGCCC TTACTGGTTT GACAATCGCT GAATACTTCC GTGATGTGGA | 5160 |
| AGGCCAAGAC GTGCTTCTCT TTATCGATAA TATCTTCCGT TCACTCAGG CTGGTTCAGA | 5220 |
| AGTATCTGCC CTTTGGGTC GTATGCCATC AGCCGTTGGT TACCAACCAA CACTTGCTAC | 5280 |
| GGAAATGGGT CAATGCAAAG AACGTATCAC ATCAACCAAG AAGGGTTCTG TAACCTCTAT | 5340 |
| CCAGGCTATC TATGTGCCAG CGGATGACTA TACTGACCCA GCGCCAGCAA CAGCCTTCGC | 5400 |
| TCACTTGGAT TCAACAACAA ACTTGAACG TAAGTTGGTA CAATTGGGTA TCTACCCAGC | 5460 |
| CGTTGACCCA CTTGCTTCAA GCTCACGTGC CTTGGCACCT GAAATCGTTG GAGAAGAGCA | 5520 |
| CTATGCAGTT GCTGCTGAAG TAAAACGTGT CCTTCAACGT TACCATGAAT TGCAAGATAT | 5580 |
| CATTGCTATC CTTGGTATGG ATGAGCTTTC TGATGAAGAA AAGACCTTGG TTGCTCGCGC | 5640 |
| CCGTCGTATC CAGTTCTTCT TGTCAAAAA CTTCAACGTT GCGGAACAAT TTACTGGTCA | 5700 |
| GCCAGGTTCT TATGTTCCAG TTGCTGAAAC TGTACGTGGC TTAAAGGAAA TCCTTGATGG | 5760 |
| TAAATACGAC CACTTGCCAG AAGATGCCTT CCGTGGTGTA GGTTCATATCG AAGATGTGAT | 5820 |
| TGCAAAAGCT GAAAAAATGG GATTTTAAGA GGTGATCTAT GGCTCAGTTA ACTGTCCAGA | 5880 |
| TCGTGACACC AGATGGTCTC GTCTATGATC ACCATGCCAG CTATGTATCG GTTCGAACTC | 5940 |
| TGGATGGTGA GATGGGGATC TTGCCACGAC ATGAAAATAT GATTGCGGTT TTAGCAGTTG | 6000 |
| ATGAAGTAAA GGTAAAACGT ATCGATGATA AAGATCACGT GAACTGGATT GCAGTAAACG | 6060 |
| GTGGCGTTAT TGAAATGCC AATGATATGA TCACAATCGT CGCTGACTCT GCAGAACGTG | 6120 |
| CTCGTGATAT CGATATCAGT CGTGCAGAAC GTGCCAACT TCGTGCAGAA CGTGCAATTG | 6180 |
| AAGAAGCACA AGACAAACAT TTGATTGACC AAGAACGTCG TGCTAAGATT GCTTTGCAAC | 6240 |

1041

| | |
|---|------|
| GTGCTATTAA CCGTATTAAT GTCGGAAATA GACTATAAGA AAAAATGAAC TTGAAAATAC | 6300 |
| CAAGTTCATT TTTTATGGTG TTTTAAGGAG CAAAACGGAT GCAGACTGCT TCGGGAACAT | 6360 |
| GGAAGTCGTT GGAGAGTCT GCTAGACGAC CATTGTCACA ATTACGTTTA AAGACAGTTG | 6420 |
| CATTGTCAGA GTCTTGATGG ACAACAATGA GAAATTTTGT GTCGGGTGTC AAATCAAAAT | 6480 |
| CACGTGGAGT CTGACCATGC GTTGAACGA TTTCTAATAA CTCTAAGCTA CCGTCCGCAA | 6540 |
| GGATGGTATA TACTGCGATA GAATCATGGC CACGGTTAGA AGCGTAGAGG TATTTACCGT | 6600 |
| CTTTAGAGAG ATGAATAGCA GCGGTTCCAT TAAAGCCTTC GTAAGCTTCC GGTAAAGTTG | 6660 |
| AAATGACCTG CATACTTCA AATTCGCCAA CGCCATCGTA GATTAAAACT TCGATAGTAC | 6720 |
| TATTGAGTTC ACAAATGAGA TAAGCGATT TATAGTGGTT ATGGAAAATG ATATGGCGTG | 6780 |
| AGCCTGCTCC TGGCTTGCTG TGATAGGTAT AGAGCTTAGA TAATTTTCCT TCTTGATCGA | 6840 |
| GGTCATAGGT GATGACTTGG TCAGTCCCA AGTCGCAGGT CACTAGATAG TGGTCAGGTG | 6900 |
| TTAAATCTGT ATAGTGAACA TGGGGGAAG CTTGATTTTC ATGTGGACCT TGGCCACTGT | 6960 |
| GTTGATCCAT ATCACTAAGT AGAAGACTAC CATCTTCCTG GCGTTTATAA ACAAGGACTT | 7020 |
| GTCCCTTGTG ATAGTTAGCT GCGTAAACCA AATCAGCCTT TTCATCGACA GCAACATAAC | 7080 |
| AGTGGGGAGC TCCTTCTTCA ACAACATGAT TTAACACAGT CCCGTCAGTT TGATAGGCTG | 7140 |
| CAATCCCCC CTTATCGTCT TGGCTACCAA CAGTGATATA ATGTTGGTGC TGGTCAAAGG | 7200 |
| CAAGGTAGGT TGGACTTGGC TCAGCTGCAA AAAGTTCTAG ATTTGAAAGC TGACCAGTTT | 7260 |
| CTGTATCAAA GTCTGCCTTG TAAATCCCTT GAGAAGTACG ACGTGTATAA GTTCCAAAAT | 7320 |
| AAACAGTTTC TTTCATTACT ATACCTCTGT GTAAAGATAA GACTATTATA TCACAAAAC | 7380 |
| AAGTAAATTA AAGATATCCA ATTAGATGTA AGCACTTTAA AAAAGAGTTA TTTTGTTC | 7440 |
| AAAATGGTAT AATGAGAGAA CAATAGAAAG GAAGTATTTA TGGAGCAAAA AGAGAAACAT | 7500 |
| TTTAGCCTAT CTTGGTTTTT CAAGTGTTTT TTAGATAACA AGGCAATTAC GGTATTTT | 7560 |
| GTAACCTTAT TATTGGGACT GAATCTTTTT ATTTTAAGTA AGATTAGTTT TCTATTTT | 7620 |
| CCTGTTTTAG ACTTTTTAGC AGTTGTGATG TTGCCAGTCA TTTGTCTGG TTTGTTATAT | 7680 |
| TATTTGTTGA ATCCTATTGT TGATTGGATG GAGAAGCATA AGGTTAATCG TGTATAGCT | 7740 |
| ATCACTATTG TCTTTGTTAT CATCGCTCTC TTTATCATTT GGGGCTTGGC AGTCGCCATT | 7800 |
| CCAAATCTGC AACGTCAGGT TTGACCTTT GCAAGAAACG TTCTGTTTA CTTAGAAGAT | 7860 |
| ATAGATAGGA TTGTTAATGG ATTGGTAGCC CAGCACCTGC CAGATGATTT CAGACCTCAA | 7920 |
| TTAGAGCAAG TTTTGACCAA TTTTCTAGC CAGGCTACAG TTTTGGCAAG TAAGGTTTCA | 7980 |

1042

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| TCTCAGGCAG | TCAACTGGGT | GAGTGCCTTT | ATTAGCGGGG | CTTCTCAAGT | GATTGTTGCC | 8040 |
| TTGATTATCG | TTCTTTTCAT | GCTCTTTTAT | CTCTTGCGTG | ATGGGAAAGG | CTTGCGTAAC | 8100 |
| TATTTGACCC | AATTCATTCC | AAGAAAATG | AAGGAACCTG | TTGGACAAGT | TTTATCAGAT | 8160 |
| GTGAATCAAC | AGTTGTCCAA | CTATGTTTCA | GGGCAAGTGA | CAGTGGCTAT | TATTGTAGCA | 8220 |
| GTAATGTTTA | TCATCTTCTT | CAAGATTATT | GGTCTACGCT | ATGCGGTTAC | GCTGGGGGTT | 8280 |
| ACTGCTGGTA | TTTTAAATCT | GGTCCCTTAT | CTTGGTAGCT | TTCTAGCCAT | GCTTCCTGCT | 8340 |
| CTAGTATTGG | GTTTGATTGC | TGGTCCAGTC | ATGCTTTTGA | AAGTAGTGAT | TGTCTTTATC | 8400 |
| GTAGAACAAA | CTATTGAAGG | CCGTTTGTG | TCTCCATTGA | TTTGGGAAG | TCAATTAAAC | 8460 |
| ATCCACCCTA | TTAATGTTCT | CTTTGTTTGT | TTAACTTCAG | GATCTATGTT | TGGTATCTGG | 8520 |
| GGAGTTTTAC | TTGGTATTCC | GGTTTATGCC | TCTGCTAAGG | TTGTCATTTC | AGCCATTTTC | 8580 |
| GAATGGTATA | AGGTAGTCAG | TGGTCTATAT | GAATTAGAGG | GTGAGGAAGT | CAAGAGTGAA | 8640 |
| CAATAGTCAA | CAGATGTTAC | AGGCTTTGGA | GGAGCAAGAT | TTAACTAAGG | CTGAGCATTG | 8700 |
| TTTCGCCAAA | GCTTTAGAAA | ATGATTCAAG | TGATCTTCTG | TATGAATTGG | CAACTTATCT | 8760 |
| TGAAGGGATT | GGTTTCTATC | CTCAGGCCAA | GGAAATTTAC | CTGAAAATTG | TAGAGGATTT | 8820 |
| TCCAGAGGTT | CATCTTAATC | TAGCTGCAAT | TGCTAGCGAG | GATGGTCAAA | TAGAAGAAGC | 8880 |
| CTTTACCTAT | CTTGAGGAAA | TCCAAGCTGA | CAGTGACTGG | TATGTCTCGT | CTTTGGCTCT | 8940 |
| GAAGGCAGAC | CTTTACCAGC | TGGAAGGTTT | GACAGATGTG | GCACGTGAGA | AATTATTGGA | 9000 |
| GGCCTTGACC | TACTCAGAGG | ATTCTCTCTT | GATATTGGGT | TTGGCAGAGT | TGGATAGTGA | 9060 |
| GTTGGAAAAT | TACCAAGCGG | CTATTCAAGC | CTATGCCCAG | TTAGATAATC | GCTCGATTTA | 9120 |
| TGAGCAAACG | GGCATTTCCA | CCTATCAACG | AATTGGCTTT | GCCTATGCTC | AGTTAGGGAA | 9180 |
| ATTTGAAACG | GCTACTGAGT | TTTTAGAAAA | AGCCCTGGAG | TTAGAATACG | ATGACTTAAC | 9240 |
| AGCTTTTGAG | TTGGCCAGTC | TTTATTTTGA | TCAAGAAGAA | TATCAAAAAG | CCACCCTCTA | 9300 |
| CTTTAAGCAG | CTTGATACCA | TTTCTCCTGA | CTTTGAAGGC | TATGAGTATG | GGTACAGTCA | 9360 |
| GGCTTTACAT | AAGGAACATC | AAGTTCAAGA | AGCCCTGCGT | ATCGCTAAGC | AAGGATTAGA | 9420 |
| GAAAAATCCC | TTTGAAACTC | GCCTCTTGCT | AGCTGCTTCA | CAATTTTCTT | ATGAATTGCA | 9480 |
| TGATGCTAGT | GGTGCAGAAA | ATTATCTCCT | TACTGCAAAA | GAAGACGCTG | AGGATACAGA | 9540 |
| AGAAATCTTG | CTTCGTTTAG | CCACTATTTA | TCTGGAGCAG | GAGCGTTATG | AGGATATTCT | 9600 |
| AGAATTGCAG | AGTGAGGAGC | CAGAAAATCT | TTTGACCAAG | TGGATGATTG | CTCGTTCTTA | 9660 |
| TCAAGAAATG | GACGATTTGG | ATACTGCTTA | TGAGTATTAT | CAAGAGTTGA | CAGGAGATTT | 9720 |
| GAAGGACAAT | CCAGAATTTT | TGGAACACTA | TATCTATCTC | TTGCGTGAAT | TGGGACATTT | 9780 |

1043

| | |
|--|-------|
| TGAAGAAGCA AAAGTCCATG CTCACACTTA CTTAAAACTG GTTCCAGATG ATGTGCAAAT | 9840 |
| GCAAGAAGCTG TTTGAGAGAT TGTAAGAATG TTTAACCCTAA ATCATTTCATA CCTCTCTCAA | 9900 |
| CTAGATGTAA CTTACAAAAC CCCTGACCTC ATGAGCCACT TTCTTCCTCC TCATGAGGTC | 9960 |
| AGTTTACTT TCTGCTGTT CAGTATCGTT TTTCTCGCT AGATTTCCTC AAAAGGGCAG | 10020 |
| ACTCCTCCCT TGGTGCGTCA CACGATTTT TCATCTCGAC TGTCTTTAA TGCATCATTA | 10080 |
| ACGACGCTT TCTTCTAGGT GGTTCATAAG GAACAGGAAG ATTCAGGTG ACTTTTCTAA | 10140 |
| TCCTAGAATA AAGTGCTGAA AACAAATTCG AATAGGCATA GAGACTAGAC AATTTGAGGA | 10200 |
| GCTGCTTGGC TCCTGTTTGA ACACATTTTC CCACCACGTG AAGAAAAAGA TGGCGGAAGC | 10260 |
| GTTTGATTGT TAAAGTTGG AAGTCACCTC CAGCTAGATG TTTGAGAAAA AGATAGAGAT | 10320 |
| TGTAGCGCAT ACAGCTCATC ATCATACGAA TTCGTTTTG ATTAAGGTTG AACTATCCGT | 10380 |
| TTTATCGCCA AAAAATCGG | 10399 |

(2) INFORMATION FOR SEQ ID NO: 161:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 9409 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 161:

| | |
|--|-----|
| GATAAGATTA AGTTAGAAAA GAAAGAACTA GGACATATCT ACCAGATTCA GGTTTTAAT | 60 |
| AGCTATGGGC AGGAAGAAAT CTATCGTGTG ATTTTGATGG AGACCAATAT TAGTTCGGTT | 120 |
| TCAACCAATA TCAAGTATGC TGCTGTCTTG ATTAATACCA GTCAGTTGGA ACAGGCTAGT | 180 |
| CAAAAGCATG AGCAATTGAT TGTGGTCGTG ATGGCTAGTT TCTGGATTTT GTCTTTACTT | 240 |
| GCCAGTCTCT ATCTAGCTAG GGTCAGTGT AGGCCCTGC TTGAGAGTAT GCAGAAGCAA | 300 |
| CAGTCTTTTG TGGAAATGC CAGTCATGAG TTACGAACTC CACTCGCAGT TTTGCAAAAT | 360 |
| CGCTTAGAGA CCCTTTTTCG TAAGCCAGAA GCTACCATTA TGGATGTGAG CGAAAGCATT | 420 |
| GCATCGAGTT TGGAAAGAAT CCGAAATATG CGTTTTTTAA CGACAAGCTT GCTGAACTTA | 480 |
| GCTCGGAGAG ATGATGGGAT TAAGCCGGAG CTTGCAGAAG TTCCAACCTAG CTTTTTTAAT | 540 |
| ACAACTTTCA CAACTACGA GATGATTGCT TCGGAAAATA ATCGTGTCTT CCGTTTTGAA | 600 |
| AATCGTATCC ATCGAACAAT TGTCACAGAT CAGCTTCTTC TGAAACAACT GATGACCATT | 660 |
| CTTTTCGATA ATGCCGTCAA GTATACTGAG GAGGATGGTG AAATTGATTT TCTTATCTCG | 720 |

1044

| | |
|--|------|
| GCGACCGATC GCAATCTTTA TTTACTTGTT TCTGATAATG GAATCGGTAT TTCGACAGAA | 780 |
| GATAAAAAGA AAATTTTGA CCGTTTTTAT CGAGTAGACA AGGCTAGAAC CCGGCAAAAA | 840 |
| GGTGGTTTGT GTTTAGGATT ATCCCTAGCC AAGCAAATTG TAGATGCTCT AAAAGGAACT | 900 |
| GTTACTGTCA AAGATAATAA ACCCAAGGGA ACAATCTTTG AAGTGAAGAT TGCCATTTCAG | 960 |
| ACACCATCTA AAAAGAAAAA ATAAAAATAT CGCTCCAATT GGGGCGATAT TTTGGATTTA | 1020 |
| TCTTCTACGT TTTCGTTTGA TAATAGACCG TTGAACTTTT AAAACAAGTA AGCTGAATCC | 1080 |
| GATTGCTGCG GCAAAGGCAA GAGCAGTTGA TAATTTTAAT GCTAAAAAGA TAAAACTAAA | 1140 |
| GATAGCAATA CAGATACAAA AAACAGCGAT ATTAATAAAA AATAGGATTT CCTTGAGATT | 1200 |
| GGCATCAGAT TCGCTTCAG GTGTATAAGC TTGGTAATGA GGAAGCTGCT GGTTTAATTC | 1260 |
| TTCTTGATAG TCTACCTCAT AGGATTGTAA TTTTCTTACG GGCATGATTC TCTCCTTAAC | 1320 |
| AGTACATACC TATTTTATCA TTTTTCGGC AGAGAATTAT TACAGAAAGG TTACAAAAAG | 1380 |
| AATAAAGTCC CTTTTCATTT TCAAAGCATG GCTGATTTTG GAGAAATGTG GTATAATTTT | 1440 |
| TCTTATGGAA AAGATTGTCA TTACAGCAAC TGCTGAAAGT ATTGAACAAG TTGAACAACT | 1500 |
| ACTCGAAGCT GCGCTAGACC GTATCTATGT CCGTGAGAAA GATTTTGGTC TTCGCTCGCC | 1560 |
| AACGACCTTT AGTTATGACC AATTACGTGA AATCGCTAAG TTGGTTCATG ATGCTGGTAA | 1620 |
| GGAATTGATC GTTGCGGTCA ATGCTCTCAT GCACCAAGAT ATGATGGACC GTATCAAGCC | 1680 |
| TTTCTTAAAC TTCTTGGAAG AAATCAAGAC AGACTATATT ACGATTGGGG ATGCAGGCGT | 1740 |
| CTTTTACGTA GTTAACCGCG ATGGTTATTC ATTTAAGACC ATCTACGATG CTTCAACCAT | 1800 |
| GGTAACTAGC AGTCGTGAGA TTAACCTCTG GGGACAAAAG GCTGGCGCAT CTGAGGCTGT | 1860 |
| TTTGGCGCGT GAAATCCAT CAGCTGAACT TTTCAAAATG CCAGAGATTT TGGAAATCC | 1920 |
| TGCTGAAGTT TTGGTTTACG GTGCTAGCGT CATCCATCAT TCTAAACGTC CACTCTTGCA | 1980 |
| AAACTACTAT AACTTTACAC ATATCGATGA TGAAAAGACG CATAAACGTG ACCTCTTCTT | 2040 |
| GGCTGAGCCA AGTGATCCAG AGAGCCACTA TTCCATTTT GAAGATAATC ATGGGACCCA | 2100 |
| TATCTTTGCC AACAATGACC TTGATTGAT GATCAAATTA ACAGAATTGG TGGAGCATGG | 2160 |
| CTTTACTCGC TGGAACTAG AAGGGCTCTA CACTCCTGGT CAGAACTTTG TTGAGATTGC | 2220 |
| AAAACCTCTT ATCCAAGCGC GTAGCTTGAT TCAAGAGGGC AACTTTAGTC ATGCTCAAGC | 2280 |
| CTTCTTGCTG GATGAAGAAG TTCGTAACT TCACCCTAAA AACCGTTTCC TTGATACAGG | 2340 |
| ATTTTATGAC TACGATCCTG ACATGGTTAG ATAAAATACA TGATTGTTG AGAGAAGGAA | 2400 |
| GATGCAACA TTTCTTCTCT CAATTTTTCG TATTTCTTCA CTATTTTACA AAAATCAGCA | 2460 |
| GGCTAGAATG CTCTATTCTA TGGGATTTT AAGAAAAGTA GTGTTCTTGA GTTTGAAAAT | 2520 |

1045

| | |
|--|------|
| TATCCTATGT TTGCAGGTGC CAAATGGCCC TTTTTTTGGT ATAATTTTTT ATAATGAAAA | 2580 |
| CGATTGGTAA TCGCTATGTT GTGGTGGATT TAGAGGCAAC TAGCACAGGT AGTAAGGCTA | 2640 |
| AAATTATCCA AGTGGGAATT GTCGTGATTG AGGACGGAGA AATCGTCGAT CACTATACGA | 2700 |
| CGGATGTCAA TCCACATGAA CCCTTGGATG CTCATATCAA AGAACTGACA GGATTGACAG | 2760 |
| ACCAACGTCT GCGCAAGCA CCTGATTTTT CGCAAGTTC CAGAAAAATA TTTGACTTGG | 2820 |
| TGGAGGATGG GATTTTGTGA GCCCATAATG TTCAGTTTGA TGCTAATCTC TTGGCGGAAA | 2880 |
| ATTTATTTTT TGAAGGCTAT GAGCTAAGAA ACCCTCGTGT TGATACGGTC GAATTGGCCC | 2940 |
| AGGTCTTTTT CCCTGAAGTG GAAAAATATA GCTTGCCGAT TTTGTGTCGA GAATTAGGAA | 3000 |
| TTCTCTTAA ACACGCACAC ACAGCCCTTT CAGATGCCCA AGCTACAGCA GAATTACTTC | 3060 |
| TTTTTTTACG GAAAAAGATG ACCCAGCTTC CTAAAGGTCT CTTGGAACGC TTGCTGGAAA | 3120 |
| TGGCTGACGC TCTCCTATAT GAGTCCTACC TGGTTATTGA GGAAACTTAT CGCAACCAAT | 3180 |
| CTATCCTGAG TTCTCCAGAC TTGGTCCAAG TTCAAGGTCT ATATTTTAAG AAAACGGAAG | 3240 |
| CTTCTCTGGA GCCACGAAAA CTATCTCAAG ACTTTTCTAA AAATATTTCT CTGTTGAACC | 3300 |
| TTGAAGTGAG GGAGGAACAA GAAAGTTTTG CTAAAGAGGT TGGCTTGCTA TTGAAAGATG | 3360 |
| AACCTGTCTC TCTGATTCAA GCGCCGACAG GGATTGGGAA AACCTATGGC TATCTCTTAC | 3420 |
| CCGCTTTATC TCAATCCAAA GAGCGACAAA TTGTTCTTAG TGTTCGACA AAGATTCTTC | 3480 |
| AAAATCAAAT CATGGAAGAA GAAGGTAAAC GCCTCAAGGA AGTGTTCCTAT ACAGATATTC | 3540 |
| ATAGCTTAAA GGGACCACAA AATTATCTGA AGTTGGATGC CTTTATCAT TCCTTGACAG | 3600 |
| AAAATGATGA AAATCGCTTA TTTAGACGCT TTAAGATGCA AGTCTTGGTC TGGCTTACTG | 3660 |
| AGACAGAGAC AGGAGATTTG GATGAAATCG GGCAACTCTA CCGTTACCAA CATTTTCTAG | 3720 |
| CAGACCTTCG TCATGATGGG AATTATCAT CCCAGAGCTT ATTTGTGACG GAAGATTTTT | 3780 |
| GGAAACGTAG TCAAGAAAGG GCAGAGACTT GCAAGCTTTT AGTGACTAAT CATGCCTATC | 3840 |
| TCGTAACCAG ACTTGAAGAT AATCCTGAAT TTGTCAGTGA CCGTTTACTG ATTATTGATG | 3900 |
| AAGTCCAAAA GATTTTGTGA GCTCTAGAAA ATCTGCTTCA AGAGACCTAC GATATACAAT | 3960 |
| CTATTATCGA TTTAATTGAT AAGGCTTTAG TAGGAGAAGA AAACAGGGTT CAACAACGGA | 4020 |
| TACTAGAAAG TATTCGCTTT GAGTGTCTCT ACTTGATAGA ACAATTTTCT TCTGGCAAAT | 4080 |
| CTAGGAAAAA TATCTTAGAT TCTCTGGACA ATCTCCATCA GTATTTTTC AATTGGAAG | 4140 |
| TAGAAGACTT TGATGAGCTG GTTCGCTATT TTACAGCTGA AGGTGATTAC TGGCTTGAAG | 4200 |
| TAACTGAAAC GAGTCAAAG AAAATTCAGA TTTCTTCTAC AAAATCAGGC CGTACTCTTC | 4260 |

1046

| | |
|--|------|
| TGTCCTCTTT ACTTCCTGAG AGTTGCCAAG TCTTGGGAGT ATCGGCTACT CTTGAGATTA | 4320 |
| GTCAGAGGGT TTCTTTGGCA GACCTTTTAG GCTATCCTGA AGCTAAATTT GTCAAGATTG | 4380 |
| AATCTCGGGG AAAACAGGAA CAAGAAGTGG TCATGGTCAA AGATTTCCCT CTGGTAACAG | 4440 |
| AAACCTCCTT AGAAGTCTAT GCCAGAGAGG TAGCTGCTTT ACTAGTGGAA ATTCAAGCTT | 4500 |
| TCCAGCAACC GATTTTGGTT CTCTTTACCG CTAAAGACAT GCTTCTAGCA GTATCGGATT | 4560 |
| TACTTACAGT TAGCCACTTG GCCCAGTATA AAAATGGGGA TGTTCATCAG CTAAAGAAAC | 4620 |
| GCTTTGAAAA AGGTGAACAA CAAATCTTGC TTGGTGACG AAGTTTCTGG GAGGGAGTTG | 4680 |
| ATTTTTC AAG CCATCCTTCT GTGATTCAAG TTGTACCGAG GCTTCCTTTC CAAAATCCTC | 4740 |
| AAGAACCCTT GACGAAAAAG ATTAATCAAG AACTGAATCA AGAAGGGAAA AATGCCTTTT | 4800 |
| ATGATTATCA ATTGCCAATG GCCATTATTC GTTTAAACA GGCTTTGGGA AGAAGTATGA | 4860 |
| GACGTGAATA CCAACGTTC TTAACCTTA TTTTGGATAG GAGAATCGTC GGAAAACGAT | 4920 |
| ACGGCAAACA AATAGTAGCA TCTCTAGCAG AAGAAGCGAC TGTAAACC ATCTCTCGAT | 4980 |
| CCGAAGTTGA CGAGGCTATT GATAGATTTT TTAATGAGCT TTGATAAATA GTATTGTATG | 5040 |
| AAAGTATAAG GTTAGTATAT ATGAAACGTT CTCTCGACTC AAGAGTCGAT TACAGTTTGC | 5100 |
| TCTTGCCAGT ATTTTCTTA CTGGTCATCG GTGTGGTGGC TATCTATATA GCCGTAGTC | 5160 |
| ATGATTATCC CAATAATATT CTGCCATTT TAGGGCAGCA GGTCGCCTGG ATTGCCTTGG | 5220 |
| GGCTTGAT TGGTTTGTG GTCATGCTCT TTAATACAGA ATTTCTTTGG AAGGTGACCC | 5280 |
| CCTTTCTATA TATTTAGGC TTGGGACTTA TGATCTTGCC GATTGTATTT TATAATCCAA | 5340 |
| GCTTAGTTGC ATCAACGGGT GCCAAAACT GGTATCAAT AAATGGAATT ACCCTATTC | 5400 |
| AACCGTCAGA ATTTATGAAG ATATCCTATA TCCTCATGTT GGCTCGTGTC ATTGTCCAAT | 5460 |
| TTACAAAGAA ACATAAGGAA TGGAGACGCA CGGTTCCGCT GGACTTTTGG TTAATTTTCT | 5520 |
| GGATGATTCT CTTTACCATT CCAGTCCTAG TTCTTTTAGC ACTTCAAAGT GACTTGGGGA | 5580 |
| CGGCTTTGGT TTTTGTAGCC ATTTTCTCAG GAATCGTTTT ATTATCAGGG GTTTCTTGGA | 5640 |
| AAATTATTAT CCCAGTATTT GTGACTGCTG TAACAGGAGT TGCTGGTTTC TTAGCTATCT | 5700 |
| TTATTAGCAA GGACGGACGA GCTTTTCTTC ACCAGATTGG AATGCCGACC TACCAAATTA | 5760 |
| ATCGGATTTT GGCTTGGCTC AATCCCTTGG AGTTTGCCCA AACAACTGACT TACCAGCAGG | 5820 |
| CTCAAGGGCA GATTGCCATT GGGAGTGGTG GCTTATTGG TCAGGGATTT AATGCTTCGA | 5880 |
| ATCTGCTTAT CCCAGTTCGA GAGTCAGATA TGATTTTAC GGTATTGCA GAAGATTTTG | 5940 |
| GCTTTATTGG CTCTGTCCTG GTTATTGCCC TCTATCTCAT GTTGATTAC CGTATGTTGA | 6000 |
| AGATTACTCT TAAATCAAAT AACCAGTTCT AACTTATAT TTCCACAGGT TTGATTATGA | 6060 |

1047

| | |
|--|------|
| TGTTGCTCTT CCACATCTTT GAGAATATCG GTGCTGTGAC TGGACTACTT CCTTTGACGG | 6120 |
| GGATTCCCTT GCCTTTCATT TCGCAAGGGG GATCAGCTAT TATCAGTAAT CTGATTGGTG | 6180 |
| TTGGTTTGCT TTTATCGATG AGTTACCAGA CTAATCTAGC TGAAGAAAAG AGCGGAAAAG | 6240 |
| TCCCATTCAA ACGGAAAAG GTTGTATTAA AACAAATTAA ATAAGGAGAA AATCATGGTA | 6300 |
| AAAGTAGCAG TTATATTAGC TCAGGGCTTT GAAGAAATTG AAGCCTTGAC AGTTGTAGAT | 6360 |
| GTCTTGCGTC GAGCCAATAT CACATGTGAT ATGGTTGGTT TTGAAGAGCA AGTAACGGGT | 6420 |
| TCGCATGCAA TCCAAGTAAG AGCAGATCAT GTCTTTGATG GAGATTTATC AGACTATGAT | 6480 |
| ATGATTGTTC TTCCTGGAGG TATGCCTGGT TCTGCACATT TACGTGATAA TCAGACCTTG | 6540 |
| ATTCAAGAAT TGCAAAGCTT CGAGCAAGAA GGAAGAAAAC TAGCAGCCAT TTGTGCGGCA | 6600 |
| CCAATTGCCC TCAATCAAGC AGAGATATTG AAAAATAAGC GATACACTTG TTATGACGGC | 6660 |
| GTTCAGAGC AAATCCTTGA TGGTCACTAC GTCAAGGAAA CAGTAGTGGT AGATGGTCAG | 6720 |
| TTGACAACCA GTCGGGTCC TTCAACAGCC CTTGCTTTG CCTACGAGTT GGTGGAGCAA | 6780 |
| CTAGGAGGGG ACGCAGAGAG TTTACGAACA GGAATGCTCT ATCGAGATGT CTTTGGTAAA | 6840 |
| AATCAGTAAA ACGGGAGTTA TTCTCTCGTT TTTTATGTGG AAAACTCAGG GAAATCATCG | 6900 |
| CTTTTTTCAT AAAAAATGC TATAATGAAG GGTATGAAAT ATCAGGATTA CATCTGGGAT | 6960 |
| TTAGGTGGAA CTTTACTGGA TAATTATGAA ACTTCAACAG CTGCATTGTG TGAAACATTG | 7020 |
| GCACTGTATG GTATCACACA AGACCATGAC AGTGCTATC AAGCTTTAAA GGTTCCTACT | 7080 |
| CCTTTTGC GA TTGAGACATT CGCTCCCAAT TTAGAGAATT TTTTAGAAAA GTACAAGGAA | 7140 |
| AATGAAGCCA GAGAGCTTGA ACACCCGATT TTATTTGAAG GAGTTTCTGA CCTATTGGAA | 7200 |
| GACATTTCAA ATCAAGGTGG CCGTCATTTT TTGGTCTCTC ATCGAAATGA TCAGGTTTGT | 7260 |
| GAAATTTTAG AAAAAACCTC TATAGCAGCT TATTTTACAG AAGTGGTGAC TTCTAGCTCA | 7320 |
| GGCTTTAAGA GAAAGCCAAA TCCCGAATCC ATGCTTTATT TAAGAGAAAA GTATCAGATT | 7380 |
| AGCTCTGGTC TTGTCATTGG TGATCGGCCG ATTGATATCG AAGCAGGTCA AGCTGCAGGA | 7440 |
| CTTGATACCC ACTTGTTTAC CAGTATCGTG AATTTAAGAC AAGTATTAGA CATATAAGAA | 7500 |
| AAAGGAATAA GATGACAGAA GAAATCAAAA ATCTGCAGGC ACAGGATTAT GATGCCAGTC | 7560 |
| AAATTCAAGT TTTAGAGGGC TTAGAGGCTG TTCGTATGCG TCCAGGGATG TACATTGGAT | 7620 |
| CAACCTCAAA AGAAGGTCTT CACCATCTAG TCTGGGAAAT TGTTGATAAC TCAATTGACG | 7680 |
| AGGCCTTGGC AGGATTTGCC AGCCATATTC AAGTTTATAT TGAGCCAGAT GATTTCGATTA | 7740 |
| CTGTTGTGGA TGATGGGCGT GGTATCCCAG TCGATATTCA GGAAAAACA GGCCGTCCTG | 7800 |

1048

| | |
|--|------|
| CTGTTGAGAC CGTCTTTACA GTCCTTCACG CTGGAGGAAA GTTCGGCGGT GGTGGATACA | 7860 |
| AGGTTTCAGG TGGTCTTCAC GGGGTGGGGT CGTCAGTAGT TAATGCCCTT TCCACTCAAT | 7920 |
| TAGACGTTCA TGTTACAAA AATGGTAAGA TTCATTACCA AGAATACCGT CGTGGTCATG | 7980 |
| TTGTCGCAGA TCTTGAAATA GTTGGAGATA CGGATAAAAC AGGAACAAC TTTCACTTCA | 8040 |
| CACCGGACCC AAAAATCTTC ACTGAAACAA CAATCTTTGA TTTTGATAAA TTAAATAAAC | 8100 |
| GGATTCAAGA GTTGGCCTTT CTAAATCGCG GTCTTCAAAT TTCAATTACA GATAAGCGCC | 8160 |
| AAGGTTTGGG ACAAACCAAG CATTATCATT ATGAAGGTGG GATTGCTAGT TACGTTGAAT | 8220 |
| ATATCAACGA GAACAAGGAT GTAATCTTTG ATACACCAAT CTATACAGAC GGTGAGATGG | 8280 |
| ATGATATCAC AGTTGAGGTA GCCATGCAGT ACACAAC TGG TTACCATGAA AATGTCATGA | 8340 |
| GTTTCGCCAA TAATATTCAT ACCCATGAAG GTGGAACACA TGAACAAGGT TTCCGTACAG | 8400 |
| CCTTGACACG TGTATCAAC GATTATGCTC GTAAAAATAA GTTACTGAAA GACAATGAAG | 8460 |
| ATAATTTAAC AGGGGAAGAT GTTCGCGAAG GCTTAACTGC AGTTATCTCA GTTAAACACC | 8520 |
| CAAATCCACA GTTTGAAGGA CAAACCAAGA CCAAAATTGGG AAATAGCGAA GTGGTCAAGA | 8580 |
| TTACCAATCG CCTCTTCAGT GAAGCTTTCT CCGATTTCCT CATGGAAAAT CCACAGATTG | 8640 |
| CCAAACGTAT CGTAGAAAA GGAATTTTGG CTGCCAAGGC TCGTGTGGCT GCCAAGCGTG | 8700 |
| CGCGTGAAGT CACACGTAAA AAATCTGGTT TGGAAATTTT CAACCTTCCA GGGAACTAG | 8760 |
| CAGACTGTTT TTCTAATAAC CCTGCTGAAA CAGAACTCTT CATCGTCGAA GGAGACTCAG | 8820 |
| CTGGTGGATC AGCCAAATCT GGTGTAACC GTGAGTTTCA GGCTATCCTT CCAATTCGCG | 8880 |
| GTAAGATTTT GAACGTTGAA AAAGCAAGTA TGGATAAGAT TCTAGCCAAC GAAGAAATTC | 8940 |
| GTAGTCTTTT CACAGCCATG GGAACAGGAT TTGGCGCAGA ATTTGATGTT TCGAAAGCCC | 9000 |
| GTTACCAAAA ACTCGTTTGT ATGACCGATG CCGATGTCGA TGGAGCCAC ATTCGTACCC | 9060 |
| TTCTTTTAAC CTTGATTAT CGTTATATGA AACCAATCCT AGAAGCTGGT TATGTTTATA | 9120 |
| TTGCCCAACC ACCAATCTAT GGTGTCAAGG TTGGAAGCGA GATTAAAGAA TATATCCAGC | 9180 |
| CGGGTGCAGA TCAAGAAATC AAATCCAAG AAGCTTTAGC CCGTTATAGT GAAGGTCGTA | 9240 |
| CCAAACCGAC TATTCAGCGT TATAAGGGGC TAGGTGAAAT GGACGATCAT CAGCTGTGGG | 9300 |
| AAACAACCAT GGATCCCGAA CATCGCTTGA TGGCTAGAGT TTCTGTAGAT GATGTGCAGA | 9360 |
| AGCAGATAAA ATCTTTGATA TGTGATGGG GATCGAGTTG TCCTCGTCG | 9409 |

(2) INFORMATION FOR SEQ ID NO: 162:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6415 base pairs
 - (B) TYPE: nucleic acid

1049

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 162:

| | |
|---|-----|
| CCTGGGAAAG TCTTGAAAAT TATGATAGAA TGGTGGAAGG AAAAATTCAG GAGAGTAGTA | 60 |
| GTGACTCAAA ATGTTGAAAG TCTTCTCGTA TCCATTGTAA TCAGTGCATA CAATGAAGAA | 120 |
| AAATATCTGC CTGGTCTAAT TGAAGACTTA AAAAATCAAA CCTATCCTAA AGAGGATATT | 180 |
| GAAATTCTAT TTATAAATGC TATGTCCACA GATGGGACCA CAGCTATCAT TCAGCAATTT | 240 |
| ATAAAGGAAG ATACAGAGTT TAACTCAATT AGATTGTATA ACAATCCTAA GAAAAATCAA | 300 |
| GCTAGTGGTT TTAACCTGGG AGTTAAACAT TCTGTAGGGG ACCTTATTTT AAAAATTGAT | 360 |
| GCTCATTCAA AAGTTACTGA GACTTTTGTA ATGAACAATG TGGCTATTAT TCAACAAGGT | 420 |

1050
CACGGATTAC TTGGCAGTTA ATCTACTCCA GACGCAAAAA AGGTAGTGGT GATGGAGAAC 1620
ACCGTCGGAC CTTCTTGATT GGTGCCGGTG ATGGTGGGGC TCTTTTATG GATAGTTACC 1680
AACATCCAAC CAGTGAATTA GAACTGGTCG GTATTTTGA TAAGGATTCT AAGAAAAAGG 1740
GTCAAAAAC TGGTGGTATT CCTGTTTGG GCTCTTATGA CAATCTGCCT GAATTAGCCA 1800
AACGCCATCA AATCGAGCGT GTCATCGTTG CGATTCCGTC GCTGGATCCG TCAGAATATG 1860
AGCGTATCTT GCAGATGTGT AATAAGCTGG GTGTCAAATG TTACAAGATG CCTAAGGTTG 1920
AAGCTCTCTT TCAGCGCGTT GAGGAGGAG GAGCGCGTTT GAGGAGGAG GAGGAGGAG

1051

| | |
|--|------|
| TTGAGCTGCC CGCAGCTAGT TFCCTAGTTT GCTCTTTGAT TTTCATTGAG TATTACTTCA | 3420 |
| TTTTCTTCTG AAATGGAATT GTTACCCAGT CTATGCTATT GAAAATACGC CAAAACCTCT | 3480 |
| AAGGGTTTGT GAGCGATATA ATCAGGTTGA TAGTTTAGTA GATCTGCTTG CTCTCCAAAT | 3540 |
| CCCCAAGTGA TGGCCAATTT CTGAATACCT GTTTCTCGAG CTCCCAGCAT ATCAAACCTG | 3600 |
| GTATCTCCGA TGATGATGGC TTGTTCTGGT GCTAGTTGAT GTGTCTGCAA GGCTTGGTGA | 3660 |
| ATGACATCTG CCTTATGGGG TGCTTCAGGG CTAGAACCAT AAATGCCATC AAAGAAATGA | 3720 |
| TGGATTTCCA AGTTTTTTC CATGTCTTGA GCAGTAGATG TATCCTTTGT CGTGGTGATG | 3780 |
| TAGAGTGGAT AACTGCTCGA TAACTCCTCA AGCAAGTCTA TAATCTGAGG AAAGAGTTGA | 3840 |
| GCTTCATAGA TGCCTTTTGC CTTATAGTAA GAACGATATA TCTGCACGGC TTCAGAAATT | 3900 |
| TGGTCTTTGG ACAGGCAGGT CGCAAACTA CTTTCGAGAG GTGGTCCCAT AAAACCACGA | 3960 |
| ATAGTTTTGG CATCAGGGCT AGGCACCCCC AGCTCTTTAA AGGTATAGGT AAAGGCATTG | 4020 |
| TGAATCCCGA TAGAACTATC AACGAGGGTT CCATCCAAAT CGAAAAAAT CGCTGTGATA | 4080 |
| GAGGTCATGG TTTCTCCTAT TTGATAAGCT TATTCTCCGA AAATTTCTTT TTGGAGCGCA | 4140 |
| CGACCAGTAG GGGTGGTAGC GAGTCCACCT TCAGCTGTTT CACGAAAGGC AGTTGGCATG | 4200 |
| CTTGCTCCTA CTTGGTACAT GGCATCGATC ACTTCATCCA CAGGGATTTT AGATTGCGATA | 4260 |
| CCTGCCAAGG CCATGTCTGC TGCAGTGAAA GCAAAGCTAG CTCCCATGGC ATTACGTTTG | 4320 |
| ACACAGGGAA CTTCGACCAA ACCTGCAACA GGGTCACAGA TGAGGCCTAG CATATTTTTA | 4380 |
| ATGACAAAGG CAATAGCTTG ACTGGCCTGA TAAGGTGTTT CACCTGCAGC CAGAGTCAAG | 4440 |
| GCGGCAGCAC TCATAGCAGA GGCTGAACCA ACTTCAGCTT GACACCCACC CTCAGCACCT | 4500 |
| GAGATGGAGG CATGTGTTGC GATGACTAGT CCAAAGGCAC CAGCAGCAAA GAGGAAATCC | 4560 |
| AATTGTTGCT CGTGGCTGAG GTCTAATTTT TCAATAGCAG CAGTGAGAAC GGATGGCAGA | 4620 |
| CAGCCAGCAC TTCCAGCGGT TGGAGTGGCA CAGACCAAGC CCATTTTGGC ATTGTGTTCA | 4680 |
| TTGACTGCGA TGGCATTTTC GGCAGCCGAG AGAATCGTAT AATCTGACAG AGTTTTTCCG | 4740 |
| TTTTCGATGT AGTGATCCAA TTTGGCAGCA TCTCCACCTG TCAGGCCACT ACGAGATTTA | 4800 |
| TTTTCATTTGA GGCCAAGTTG GACAGAGGCT TTCATAACTT CCAGATTGCG TTCCATGAGA | 4860 |
| AGGAAGACTT CTTACGTTT GCGACCGGTC AATTCAAACCT CTGTTGTAAT CATGAGTTCT | 4920 |
| GCGACATTTT CTTGAAAGTC CAGATCTGCT TGCTCGACCA ATTCTTTGAT AGAATAAAAC | 4980 |
| ATGCTTCCTC CTATTTAAAG AAATGACAT TGTGGAGATG AGGGATTTT CGAATTTCTT | 5040 |
| CGATAGCCTC ATCACAGTTG CGACTGTCAA CTTGATAAT CATAATGGCT TTTTCACCAG | 5100 |

1052

| | |
|--|------|
| CTTTTTCACG AGTGACATTC ATCTGGGCGA TATTGATACC ATAGCGGGAA AGCGCCTCTG | 5160 |
| TAACAAGGGC AATCATACCT GGAATATCTT GATGAACGAT GATGATAGTC GGTGTATTCA | 5220 |
| TATTGAGAGA GACGGCAAAA CCATTGAGTT CGGTTACCTG AATATTTCCCT CCACCGATAG | 5280 |
| AAATACCAGT CACGCTGATG GTCTTGTGGG CATTTTAAAC AGTAATTTTA GTGGTGTTAG | 5340 |
| GGTGAGGGGC ATTGCTGTCT TTCTGAATGG TCCAGACAAT CTTGATACCA CGCTTGTGGG | 5400 |
| CAATTTCCAG ACTATTTGGA ATTTTCAGGAT CATCTGTATC CATTCCTAAA ATACCTGCAA | 5460 |
| CAAGGGCTAG GTCTGTTCCG TGACCACGAT AGGTCTTGGC AAATGAGTTA AAAAGTTGGA | 5520 |
| ATTCAACTTC TGTCCGAGTA TCATCAAAAA TGGAAGAGAC AATCTTCCCA ATACGAACAG | 5580 |
| CACCAGCGGT ATGGCTACTA GATGGGCCAA TCATAACTGG TCCGATGATA TCAAAGACAG | 5640 |
| ATTGAAAACG AAGTGATTTT ATCAGTTTCC CCTTATAAAA ATTCTTATCT CTATTATATC | 5700 |
| AAAGAATGAG GGGCTTGGCT TTAATTGTGG ATGAAAACCT TTCTAATACC TCAAATAGCA | 5760 |
| TAAAAATAGT ATCTTTTATG ACAAAAAACA CCTTATTTAG GGAAATAAAA AATAATTTTG | 5820 |
| TAATATTTCT ACATAAAAGT GTCAAGAAAC GGTAATATTT AAAGGGTATG ATAGAACTAT | 5880 |
| AGAAAGAAGG AGAATTTTCG AATATGAAAT CAATAACTAA AAAGATTAAA GCAACTCTTG | 5940 |
| CAGGAGTAGC TGCCTTGTTT GCAGTATTTG CTCCATCATT TGTATCTGCT CAAGAATCAT | 6000 |
| CAACTTACAC TGTTAAAGAA GGTGATACAC TTTCAGAAAT CGCTGAAACT CACAACACAA | 6060 |
| CAGTTGAAAA ATTGGCAGAA AACAACCACA TTGATAACAT TCATTTGATT TATGTTGATC | 6120 |
| AAGAGTTGGT TATCGATGGC CCTGTAGCGC CTGTTGCAAC ACCAGCGCCA GCTACTTATG | 6180 |
| CGGCACCAGC CGCTCAAGAT GAAACTGTTT CAGCTCCAGT AGCAGAAACT CCAGTAGTAA | 6240 |
| GTGAAACAGT TGTTTCAACT GTAAGCGGAT CTGAAGCAGA AGCCAAAGAA TGGATCGCTC | 6300 |
| AAAAAGAATC AGGTGGTAGT ATACAGCTAC AAATGGACGT TATATCGGAC GTTACCAATT | 6360 |
| AACAGATTCA TACCTGAACG GTGACTACTC AGCTGAAAAA CAAGAACGGG TACCG | 6415 |

(2) INFORMATION FOR SEQ ID NO: 163:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 8494 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 163:

| | |
|---|-----|
| TACCCCTTTC GAATTTTGGC AAAAATTCGG TAAGGCTTTG ATGGTAGTTA TCGCGGTAT | 60 |
| GCCGGCTGCT GGTTTGATGA TTTCAATCGG TAAGTCTATC GTGATGATTA ACCCAACCTT | 120 |

1053

| | |
|--|------|
| TGCACCACTT GTCATCACAG GTGGAATTCT TGAGCAAATC GGTGGGGGG TTATCGGTAA | 180 |
| CCTTCACATT TTGTTTGCCC TAGCCATTGG AGGAAGCTGG GCTAAAGAAC GTGCTGGTGG | 240 |
| TGCTTTCGCC GCTGGTCTTG CCTTCATCTT GATTAAACCGT ATCACTGGTA CAATCTTTGG | 300 |
| TGTATCAGGC GATATGTTGA AAAATCCAGA TGCTATGGTA ACTACTTTCT TTGGTGGTTC | 360 |
| AATCAAAGTT GCTGATTACT TTATCAGTGT TCTTGAAGCT CCAGCCTTGA ACATGGGGGT | 420 |
| ATTTCGTAGG ATTATCTCAG GTTTTGTAGG GGCAACTGCT TACAACAAAT ACTACAACCT | 480 |
| CCGTAAACTT CCTGATGCAC TTTCATTCTT CAACGGGAAA CGTTTCGTAC CATTTGTAGT | 540 |
| TATTCTTCGT TCAGCAATCG CTGCAATTCT ACTTGCTGCT TTCTGGCCAG TAGTTCAAAC | 600 |
| AGGTATCAAT AACTTCGGTA TCTGGATTGC CAACTCACAA GAACTGCTC CAATCTTGC | 660 |
| ACCATTCTTG TATGGTACTT TGAACGTTT GCTCTTGCCA TTGGTCTTC ACCACATGTT | 720 |
| GAATATCCCA ATGAATACA CAGCTCTTGG TGGTACTTAT GACATTTTAA CTGGTGCAGC | 780 |
| TAAAGGTACT CAAGTATTCG GTCAAGACCC ACTATGGCTT GCATGGGTAA CAGACCTTGT | 840 |
| AAACCTTAAA GGTACTGATG CTAGTCAATA TCAACACTTG TTAGATACAG TACATCCAGC | 900 |
| TCGTTTCAAA GTTGGACAAA TGATCGGTTC ATTCGGTATC TTGATGGGTG TGATTGTTGC | 960 |
| TATCTACCGT AATGTTGATG CTGACAAGAA ACATAAATAC AAAGGTATGA TGATTGCAAC | 1020 |
| AGCTCTTGCA ACATTCTTGA CAGGGGTAC TGAACCAATC GAATACATGT TCATGTTTAT | 1080 |
| CGCAACACCT ATGTATCTTG TTTACTCACT TGTTCAAGGT GCTGCCTTCG CTATGGCTGA | 1140 |
| CGTCGTAAAC CTACGTATGC ACTCATTGCG TTCAATCGAG TTCTTGACTC GTACACCTAT | 1200 |
| TGCAATCAGT GCTGGTATTG GTATGGATAT CGTTAACTTC GTTTGGGTAA CTGTTCTCTT | 1260 |
| TGCTGTAATC ATGTACTTTA TCGCAAACCT CATGATTCAA AAATTCAACT ACGCAACTCC | 1320 |
| AGGGCGCAAC GGAAACTACG AAAGTCTGA AGGTTTCAGAA GAAACCAGCA GCGAAGTGAA | 1380 |
| AGTTGCAGCA GGCTCTCAAG CTGTAAACAT TATCAACCTT CTTGGTGGAC GTGTAAACAT | 1440 |
| CGTTGATGTT GATGCATGTA TGACTCGTCT TCGTGTAAT GTTAAAGATG CAGATAAAGT | 1500 |
| AGGAAATGCA GAGCAATGGA AAGCAGAAGG AGCTATGGGT CTTGTCATGA AAGGACAAGG | 1560 |
| GGTTCAAGCT ATCTACGGTC CAAAAGCTGA CATTTTGAAA TCTGATATCC AAGATATCCT | 1620 |
| TGATTTCAGT GAAATCATTC CTGAAACTCT TCCAAGCCAA ATGACTGAAG CACAACAAAA | 1680 |
| CACTGTTCAc TTCAAAGATC TTACTGAGGA AGTTTACTCA GTAGCAGACG GTCAAGTTGT | 1740 |
| TGCTTTGGAA CAAGTAAAGG ATCCAGTATT TGCTCAAAAA ATGATGGGTG ATGGATTTGC | 1800 |
| AGTAGAACCT GCAAATGGAA ACATTGTATC TCCAGTTTCA GGTACTGTGT CAAGCATCTT | 1860 |

| | | | | |
|------------|-------------|------------|------------|-----------------------------|
| 1054 | | | | |
| CCCAACAAAA | CATGCTTTTG | GTATTGTGAC | GGAAGCAGGT | CTTGAAGTAT TGGTTCACAT 1920 |
| TGGTTTGGAC | ACAGTAAGTC | TTGAAGGTAA | ACCATTTACA | GTTTCATGTTG CTGAAGGACA 1980 |
| AAAAGTTGCA | GCAGGAGATC | TCCTTGTAC | AGCTGACTTG | GATGCTATCC GTGCAGCAGG 2040 |
| ACGTGAAACT | TCAACAGTAG | TTGTCTTCAC | AAATGGTGAT | GCAATTAAAT CAGTTAAGTT 2100 |
| AGAAAAACA | GGTCTCTTG | CAGCTAAAAC | AGCAGTTGCT | AAAGTAGAAT TGTAAATATAC 2160 |
| TTGAGGTTGG | AAGCTGTATT | CCAACCTCTT | ATTTTGGGAG | AAAAGAATGA AATTTTAAAC 2220 |
| ACTCAATACT | CACAGTTGGA | TGGAGAAAGA | AGCAGAGGAA | AAATTCAGG TTTTGCTTGA 2280 |
| AGATATCTCT | GAAAAGGACT | ATGATTGAT | TTGTTTTCAA | GAAATCAATC AGGAGATGAC 2340 |
| CTCGTCAGAG | GTGGAGGTTA | ATGACCTTTA | TCAAGCTTTG | CCAGCAGCTG AGCCTATTCA 2400 |
| CCAAGACCAT | TATGTTAGAC | TCTTGGTTGA | AAAGTTGTCT | GAGCAAGGGA AAAATTACTA 2460 |
| CTGGACCTGG | GCCTATAACC | ATATCGGCTA | TAACCGCTAC | CACGAAGGTG TGGCTATCTT 2520 |
| GTCTAAACA | CCTATTGAAG | CCAGAGAAAT | TTTGGTTTCA | GATGTGGATG ATCCAACAGA 2580 |
| CTATCATACT | CGCCGTGTG | CCCTAGCTGA | AACTGTAGTC | GATGGCAAGG AGCTAGCAGT 2640 |
| TGCCAGTGTT | CATCTCTCTT | GGTGGGATAA | AGGTTTCCAA | GAAGAATGGG CACGATTTGA 2700 |
| GGCTGTCTTG | AAAAAATTGA | ACAAGCCACT | TTTACTAGCT | GGAGATTTCA ACAATCCGGC 2760 |
| TGGACAGGAA | GGTTACCAAG | CTATTTTAGC | TAGTCCATTA | GGCTTACAAG ACGCATTTGA 2820 |
| AGTTGCTCAA | GAGAAAAGTG | GTAGCTATAC | TGTTCCGCCT | GAAATTGATG GCTGGAAAGG 2880 |
| GAACACTGAA | CCCCTTCGAA | TGATTATGT | CTTACTACC | AAAGAGTTAG CGGTGGAAAA 2940 |
| TTTACATGTC | GTATTTGATG | GTAACAAGAG | TCCACAAGTG | AGTGATCACT ATGGCTTGAA 3000 |
| TGCTATATTA | AACTGGAAAT | AATAACTGAA | AAGAGGTTGG | AACTATAAAA TTCCAGCCTT 3060 |
| TTCTTACTAG | AGAAGCTACT | GGAAATAGCC | TAAATAAGTG | AGACTACTGT AATGGAATAA 3120 |
| AATATGGTAT | AATTGATAAG | GTAGATAGAA | TCGAGGATGT | TATGTCAITTT ACGAAATTTT 3180 |
| AATTTAAAAA | CTATATTAGA | GAAGCCTTGA | AGGAGTTAAA | ATTTACAACCT CCAACAGAGG 3240 |
| TGCAAGACAA | GTTGATTCCCT | ATTGTTTTGG | CAGGTCGTGA | CCTAGTAGGA GAATCAAAAA 3300 |
| CAGGTTCAAG | TAAGACTCAT | ACTTTCTTGT | TACCGATTTT | CCAGCAATTA GATGAAGCTA 3360 |
| GCGATAGTGT | ACAAGCAGTG | ATTACTGCAC | CGAGTCGTGA | GTTGGCTACT CAAATTTACC 3420 |
| AAGTAGCGCG | TCAGATTTCA | GCTCACTCAG | ATGTCGAAGT | TCGTGTGGTT AATTATGTGG 3480 |
| GTGGTACGGA | TAAGGCTCGC | CAGATTGAGA | AATTGGCAAG | CAATCAGCCT CATATTGTTA 3540 |
| TTGGAACACC | AGGCCGTATC | TACGACTTGG | TTAAATCTGG | TGATTTAGCT ATTCATAAAG 3600 |
| CCAAGACATT | TGTTGTTGAT | GAAGCAGATA | TGACCTTGGA | TATGGGATTC TTGGAAACTG 3660 |

1055

| | |
|--|------|
| TTGATAAGAT TGCTGGCAGT CTTCCAAAAG ACTTGCAATT CATGGTCTTC TCAGCGACTA | 3720 |
| TCCCACAAAA ACTGCAACCA TTCTTGAAAA AATACTTATC AAATCCTGTT ATGGAGAAAA | 3780 |
| TTAAGACCAA AACGGTTATT TCTGACACCA TTGATAATTG GTTGATTTTCG ACCAAGGGAC | 3840 |
| ATGATAAGAA TGCTCAAATT TACCAGTTGA CTCAGTTGAT GCAGCCGTAT TTGGCAATGA | 3900 |
| TTTTTGTTAA CACTAAAACG CGTGCTGATG AATTGCATTC ATATCTGACT GCTCAAGGCT | 3960 |
| TGAAGGTTGC AAAAATCCAT GGCATATTG CCCCTCGTGA ACGCAAGCGA ATCATGAATC | 4020 |
| AGGTGCAAAA TCTGGATTTT GAGTATATTG TCGCAACAGA TTTGGCAGCG CGTGGGATTG | 4080 |
| ACATTGAAGG TGTCAGCCAT GTCATCAATG ATGCCATTCC GCAAGACTTA TCTTTTTTTG | 4140 |
| TTCATCGTGT TGGTCGFACT GGACGAAATG GCCTACCAGG TACAGCTATT ACCCTTTTATC | 4200 |
| AGCCAAGTGA TGA CTGAT ATCCGTGAGT TGGAGAAATT GGAATCAAG TTAGTCCTA | 4260 |
| AGATGGTCAA AGACGGGGAA TTTCAAGATA CCTATGACCG TGATCGTCGT GCCAACCCTG | 4320 |
| AGAAAAACA AGATAAACTT GATATCGAAA TGATTGGTTT GGTAAAAAG AAAAAGAAAA | 4380 |
| AAGTCAAACC GGGTTATAAG AAGAAAATC AATGGGCGGT TGATGAAAAG CGCCGTAAAA | 4440 |
| CCAAGCGTGC TGA AATCGC GCTCGCGGTC GTGCAGAGCG TAAAGCTAAA CGCCAAACAT | 4500 |
| TTTAATAGAA ATGTTGGAG TATTGAGCTC CACTTTTTT ATTTATGAGA ACGAACTATC | 4560 |
| TAAACCGAAA CACTACATTA AAGACTGCAA ATTGCGATTA AAAATGGTAT AATGATAAAG | 4620 |
| TTATATAGTC CCGATAAGAT GGTAGGTATT TATTACGAAG AGTTTTCTTA TCAGTACTTT | 4680 |
| GTAACCTAT AACAAATTTT TTTAAGGGGG GACATTTTTA TGTCAGAGCG TAAATTATTC | 4740 |
| ACGTCTGAAT CTGTATCTGA GGGGCATCCG GATAAGATTG CAGACCAAAT TTCAGATGCG | 4800 |
| ATTTTGGATG CTATTTTAGC AAAGGATCCA GAGGCGCACG TTGCTGCTGA AACAGCTGTA | 4860 |
| TATACTGGTT CTGTCCACGT TTTTGGTGAA ATTTCTACAA ATGCCTATGT GGATATTAAC | 4920 |
| CGTGTGGTTC GTGATACCAT TGCAGAGATT GGTATACCA ATACAGAATA TGGATTTTCT | 4980 |
| GCTGAGACGG TGGGAGTACA CCCATCTTTG GTGGAACAAT CTCCTGACAT CGCTCAAGGT | 5040 |
| GTTAACGAAG CCTTGGAGGT TCGTGGAAT GCTGATCAAG ATCCACTGGA CTTGATTGGA | 5100 |
| GCAGGTGACC AAGGGCTCAT GTTTGGATTT GCAGTAGATG AAACAGAAGA GCTTATGCCA | 5160 |
| TTGCCAATTG CACTCAGTCA TAAATTGGTT CGTCGTCTGG CAGAACTTCG TAAGTCTGGA | 5220 |
| GAAATTAGCT ATCTCCGTCC AGATGCAAAA TCACAAGTTA CAGTTGAGTA CGATGAAAAT | 5280 |
| GACCGTCCGG TACGTGTAGA TACAGTCGTT ATTTCTACTC AGCATGATCC AGAGGCCACT | 5340 |
| AATGAACAAA TCCATCAAGA TGTGATTGAC AAGGTCATCA AAGAAGTTAT TCCATCTTCT | 5400 |

1056

| | |
|--|------|
| TATCTTGATG ATAAGACAAA ATTCTTTATC AATCCGACAG GTCGTTTGT AATCGGTGGT | 5460 |
| CCTCAAGGGG ACTCAGGTTT GACTGGTCGT AAGATTATTG TAGATACTTA TGGTGGCTAC | 5520 |
| TCTCGTCATG GTGGTGGTGC CTTCTCTGGT AAAGATGCCA CTAAGGTGGA TCGTTCAGCC | 5580 |
| TCTTATGCGG CTCGCTATAT TGCCAAGAAT ATCGTTGCAG CAGACCTTGC TAAGAAGGCA | 5640 |
| GAAGTGCAGT TGGCCTATGC TATCGGTGTT GCGCAACCTG TTTCTGTTCG TATCGATACT | 5700 |
| TTTCGGTACAG GAACAGTAGC TGAAAGTCAA CTTGAAAAAG CGGCTCGTCA AATCTTTGAC | 5760 |
| CTTCGCCCTG CAGGGATTAT CCAAATGCTG GACCTCAAGC GTCCAATTTA CCGTCAAACA | 5820 |
| TCGGCTTACG GTCACATGGG ACGTACAGAT ATTGATCTTC CATGGGAACG TTTGGATAAG | 5880 |
| GTAGATGCTT TGAAAGAAGC AGTAAAATAA GATTTTAAGA GGGGAACGTC CTCTCTTTT | 5940 |
| TATAGTTTTT AACTATACTG GGATACTGTT CTGAAAATCC ATTTTGCGAA AGTAGAGATT | 6000 |
| TACATGTATA GTAGATTGAA ACTAGAATAG TACACCTCAA CTTCTAAAAC ATTGTTAGCA | 6060 |
| ATCAATTTGA CTGTCCTGAT CGATTCTCC TGTCTTGTTC TCATTTTACT ATATTTCTTT | 6120 |
| AAAAATGATA AAGGTTAAGA TTTCTCCTCG TAATAGATAA TCTTGGGGAT ATTTCAATCC | 6180 |
| AAAGTTTTAT TCGTTATCAC TTGACTATTG CAAGGTTTTT TAGAGCAACA GAGTCATGGA | 6240 |
| ATGGACTCAT GGTGAGATT TCTCCTTGMT GCTTGGACTT CATTCAAAAG TCTGTTACCC | 6300 |
| AAGCCTTGTT CAAACTTCTA ATACACTAGC TGTTTCCATA GCATGACTTC TGTACTAGAC | 6360 |
| TTTCTTTTCC GAATAAATAG ATAGAACCAC AGAATCTAGT AAACCTAGAA TTAAAATTAT | 6420 |
| GGTATAATAT TAGCAATAAA AGAAATCTGG AGGATTAGAA TCATGGTATC AACGAAAACA | 6480 |
| CAAATTGCTG GTTTTGAGTT TGACAATTGC TTGATGAATG CAGCAGGTGT GGCTTGATG | 6540 |
| ACGATAGAGG AGTTAGAAGA GGTCAAAAAC TCAGCGGCAG GAACCTTTGT TACTAAGACA | 6600 |
| GCGACCTTGG ACTTCCGTCA GGGGAATCCT GAGCCACGCT ACCAAGATGT TCCACTTGGT | 6660 |
| TCCATCAACT CTATGGGCTT GCCAAATAAT GGCTTAGACT ATTATTTGGA TTATCTTTTA | 6720 |
| GATTTGCAGG AAAAAGAGTC GAACCGAACT TTCTTCTTAT CTCTGGTCGG CATGTCCTCA | 6780 |
| GAGGAAACCC ATACTATTTT GAAAAAGTC CAAGAGAGTG ATTTTCGTGG TCTGACTGAG | 6840 |
| CTAAATCTTT CCTGTCCAAA TGTTCAGGT AAACCTCAGA TTGCCTATGA TTTTGAGACA | 6900 |
| ACAGACCGGA TTTTGGCAGA AGTGTGCT TACTTCACCA AACCTCTGG AATTAAATTG | 6960 |
| CCACCTTATT TTGATATTGT TCACTTTGAC CAAGCGGCAG CTATTTTCAA CAAATATCCG | 7020 |
| CTCAAGTTTG TCAACTGCGT TAACTCTATC GGAAACGGCC TCTATATAGA AGACGAATCT | 7080 |
| GTCGTTATTC GGCCTAAGAA TGGTTTTGGT GGAATTGGTG GAGAATACAT CAAACCGACT | 7140 |
| GCTTTAGCCA ATGTTACGCG CTTTTATCAA CGTTTAAATC CTCAAATCCA AATTATCGGA | 7200 |

1057

| | |
|---|------|
| ACAGGTGGCG TTCTGACTGG TCGAGATGCC TTTGAACACA TCCTCTGTGG AGCAAGTATG | 7260 |
| GTGCAGGTGG GAACGACCCT TCACAAAGAA GGCCTCAGTG CTTTGTACCG CATTACCAAT | 7320 |
| GAAC TGAAAG CAATCATGGT GGAAAAAGGC TACGAGAGCT TAGAAGATT T CCGTGGGAAA | 7380 |
| TTGCGCTATA TTGACTAAAT TAAATCGAAA AATCTGAAGA AAGGAGAGAC GATGCTAGCC | 7440 |
| ATTGAAGAAA GTCAGAAGTT GACTTTATCA AATTTACCGA GCCTGAGCCT ATTTACAGGG | 7500 |
| ACAGATCAGG GTCAGTTTGA AGTGATGAAG AGTCAAATGT TGAAACAGAT TGGGTATGAT | 7560 |
| TCTGCTGACC TCAACTTTGC CTACTTTGAT ATGAAAGAAG TAGTTTACAA GGATGTGGAA | 7620 |
| CTGGAGTTGG TCAGCCTTCC TTTCTTTGCG GATGAGAAAA TCGTGATATT AGATTATTTT | 7680 |
| ATGGATATCA CGACTGCTAA GAAACGCTTT TTGACAGATG ATGAGCTTAA GTCATTTGAG | 7740 |
| GAATACCTTG ACAATCCTTC TCCAACAACC AAGTTGATAA TCTTTGCAGA AGGAAAGCTG | 7800 |
| GATAGCAAAA GACGGTTAGT CAAATTACTT AAGCGTGATG CCAAGGCCTT CGATGCAGTA | 7860 |
| GAAGTAAAAG AACAAGAATT GCGCCAGTAC TTCCAAAAGT GGAGTCAGAA ACAAGGTCTG | 7920 |
| CAGTTTACCA ATCATTCTTT TGAAATCTC CTCATCAAGT CGGGGTTTCA ATTTAGCGAA | 7980 |
| ATCCAGAAAA ATCTTCTCTT TTTACAGTCC TATAAGGCGA ATTCTGTTAT TGAGGAAGAG | 8040 |
| GATATTGTTA ACGCAATTCC CAAGACTTGC AGGACAATAT TTTTGATTTA ACTCAGTTTA | 8100 |
| TTCTGACTAA AAAGATGGAT CAGGCGCGCG ATTTGGTGAG AGACTTGACC TTGCAAGGGG | 8160 |
| AAGATGAAAT CAAACTGATT GCAGTCATGC TGGGACAATT TCGGACTTTT ACTCAGGTGA | 8220 |
| AGATTTTGGC GGAGTCTGGC CAAACAGAAT CGCAGATTGC AAGTAGTTTA GGTAGTTATC | 8280 |
| TGGGACGTAA CCCAAATCCT TATCAAATCA AGTTTGCAAT AAGAGATTCG AGAGGACTTT | 8340 |
| CTTTGAGCTT TTTGAAGCAA GCTATTTCTT ATTTGATTGA GACAGACTAT CAGATTAAGA | 8400 |
| CAGGTCTTTA TGAAAAAGGT TTCCTTTTGT AAAAGGCACT CTTACAGATT GCTAGTCAGG | 8460 |
| TCAATTGACA TTTGTGAAA CTACTAACCC GCGG | 8494 |

(2) INFORMATION FOR SEQ ID NO: 164:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 9707 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 164:

| | |
|---|----|
| CCGGTCAGTT CGTTCAGTAC AAGGAATCAT AATGAACGAT CAATCAGAAA AAAAGACTAG | 60 |
|---|----|

| | | | | |
|---|------------|-------------|------|--|
| 1058 | | | | |
| AAAGAAGACT GTATGGATAA TCGACCAATT GGTTTTGG | ATTCGGGTGT | CGGGGCTTG | 120 | |
| ACCGTTGTGC GCGAGCTCAT GCGCCAGCTT CCCCATGAAG | AAATCGTCTA | TATTGGAGAT | 180 | |
| TCGGCGCGGG CGCCCTATGG CCCCCGTCCT GCTGAGCAAA | TTCGTGAATA | TACTTGGCAG | 240 | |
| CTGGTCAACT TTCTCTTGAC CAAGGATGTC AAAATGATTG | TCATTGCTTG | TAACACTGCG | 300 | |
| ACTGCGGTGCTG TCTGGGAAGA AATCAAGGCT CAACTAGATA | TTCTGTCTT | GGGTGTAATT | 360 | |
| TTGCCAGGAG CTTGGGCAGC CATCAAGTCC AGTCAAGGTG | GGAAAATCGG | AGTGATTGGA | 420 | |
| ACGCCCATGA CGGTACAATC AGACATATAC CGTCAGAAAA | TCCATGATCT | GGATCCCGAC | 480 | |
| TTACAGGTGG AGAGCTTGGC CTGTCCCAAG TTTGCTCCCT | TGGTTGAGTC | AGGTGCCCTG | 540 | |
| TCAACCAAGTG TTACCAAGAA GGTGGTCTAT GAAACCCTGC | GTCCCTTGGT | TGGAAAGGTG | 600 | |
| GATAGCCTGA TTTTGGGCTG TACTCATTAT CCACTCCTTC | GCCCTATTAT | CCAAAATGTG | 660 | |
| ATGGGGCCAA AGGTCAGCT CATCGATAGT GGGGCAGAGT | GCGTACGGGA | TATCTCAGTC | 720 | |
| TTACTCAATT ATTTTGAAAT CAATCGTGGT CGCGATGCTG | GACCACTCCA | TCACCGTTT | 780 | |
| TACACAACAG CCAGTAGCCA AAGTTTGTCA CAAATTGGTG | AAGAATGGCT | GGAAAAAGAG | 840 | |
| ATTCATGTGG AGCATGTAGA ATTATGACAA ATAAAAATTA | TGAATATAAG | GATGACCAGG | 900 | |
| ACTGGTATGT TGGGTCTTAT AGTATTTTGT GTGGCGTTAA | CAGTTTGAGC | GACTATAAGA | 960 | |
| CAGATTTTCC TCTGTTTGA TTTCTCCAAA TATTGGAGA | TGAAGAGTAT | GGTTTCCCGC | 1020 | |
| TTTCAGTTAC TGTTTTACGC TATGGTTCTA TCTACCGTTT | GTTCTCCTTT | GTGGTAGACA | 1080 | |
| TGCTTAATCA AGAAATGGGA CGAACTTGG AAGTTATTCA | ACGTCATGGG | GCCCTGCTCT | 1140 | |
| TGGTTGAAAA TGGGCAACTC TTGTATGTAG AATGCGCTAA | AGAAGGGGTC | AATGTTTCATG | 1200 | |
| ATTTCTTTGA GACAAGCAAG GTCAGAGAAA CCTGTGTGAT | TGCGACTCGT | AACGAAGGTA | 1260 | |
| AAACCAAGGA ATTCCGAGCT ATCTTTGATA AGTTAGGCTA | CGATGTGGAA | AATCTTAATG | 1320 | |
| ACTACCCTGA CCTGCCTGAA GTAGCAGAAA CAGGTATGAC | CTTTGAAGAA | AATGCCCGCC | 1380 | |
| TTAAGGCAGA AACCATTTCT CAATTAACGG GCAAGATGGT | TTTGGCAGAT | GATTCTGGTC | 1440 | |
| TCAAAGTCGA TGTCTTGGT GGCTTACCAG GCGTCTGGTC | AGCTCGTTTC | GCAGGTGTGG | 1500 | |
| GAGCAACTGA CCGTGAAAAT AATGCCAAAC TCTTGACGGA | ATTGGCCATG | GTCTTTGAAC | 1560 | |
| TCAAGGACCG CTCGGCTCAG TTCCACACAA CCTAGTCGT | AGCCAGCCCA | AATAAGGAAA | 1620 | |
| GTTTAGTTGT TGAAGCAGAC TGGTCAGGTT ATATTAACTT | TGAACCTAAG | GGTGAAAATG | 1680 | |
| GCTTTGGCTA TGATCCCCTC TTCCTGTAG GAGAAACAGG | TGAGTCATCA | GCTGAATTAA | 1740 | |
| CCCTGGAAGA AAAAAATAGT CAATCTCACC GTGCCTTAGC | CGTTAAGAAA | CTTTTGGAGG | 1800 | |
| TATTTCCATC ATGGCAAAGC AAACCATCAT TGTAATGAGC | GATTCCCATG | GCGATAGCTT | 1860 | |

1059

| | |
|---|------|
| GATTGTGGAA GAAGTCCGTG ATCGCTATGT GGGCAAAGTC GATGCTGTTT TTCATAACGG | 1920 |
| CGATTCTGAA CTACGTCCGG ATTCTCCACT TTGGGAGGGC ATCCGCGTTG TTAAAGGGAA | 1980 |
| CATGGACTTC TACGCCGGCT ACCCAGAACG TCTGGTGACT GAGCTTGGTT CGACCAAGAT | 2040 |
| TATCCAAACT CATGGTCACT TGTTTGACAT CAATTTCAAC TTTCAAAAGT TGGACTACTG | 2100 |
| GGCTCAGGAG GAAGAGGCCG CTATCTGCCT CTATGGTCAC TTGCATGTGC CAAGTGCTTG | 2160 |
| GTTGGAAGGC AAGATCCTCT TTCTAAATCC AGGTTCTATC AGTCAACCAC GAGGTACCAT | 2220 |
| CAGAGAATGT CTCTATGCTC GTGTGGAGAT TGATGATAGT TACTTCAAAG TGGACTTTTT | 2280 |
| GACACGAGAT CACGAGGTGT ATCCAGGTTT GTCCAAGGAG TTAGCCGAT GATTGCCAAG | 2340 |
| GAGTTTGAGA CTTTCTTGTT GGGCAGGAG GAACTTTTT TGACCCCTGC TAAAAATCTA | 2400 |
| GCTGTGTGA TTGATACCCA CAATGCGGAT CATGCGACCC TCTTGCTCAG TCAGATGACC | 2460 |
| TATACCCGTG TTCCCGTTGT GACAGATGAA AAACAGTTTG TTGGGACGAT TGGACTCAGA | 2520 |
| GATATTATGG CTTATCAGAT GGAGCATGAC TTGAGCCAAG AAATCATGGC GGATACGGAT | 2580 |
| ATCGTTTATA TGACAAAAAC GGACGTAGCG GTTGTTTTCG CTGATTTTAC CATTACGGAG | 2640 |
| GTCTTGACA AGCTAGTAGA TGAGTCCTT TACCGGTTG TGGATGCAGA GGGTATTTTC | 2700 |
| CAAGGGATTA TTACGCGCAA GTCCATCCTC AAGGCCGTTA ATGCCCTCTT GCATGACTTT | 2760 |
| AGTAAGGAAT ATGAGATTCG ATGCCAATGA GAGACAGGAT TTCAGCCTTT TTAGAGGAAA | 2820 |
| AGCAGGGCTT GTCTGTCAAT TCCAAGCAGT CCTATAAGTA TGATTGGAG CAATTTTATAG | 2880 |
| ACATGGTAGG TGAGCGGATT TCTGAGACCA GTCTCAAGAT TTACCAAGCC CAGCTAGCCA | 2940 |
| ATCTAAAAAT CAGCGCCAG AAGCGAAAGA TTTCGGCCTG TAACCAATTT CTATACTTTC | 3000 |
| TCTATCAAAA AGGAGAGGTG GACAGCTTTT ACCGCTTGA ATTAGCCAAA CAAGCTGAAA | 3060 |
| AGAAGACGGA AAAGCCAGAG ATTCTATACC TAGACTCTTT TTGGCAGGAA AGCGACCATC | 3120 |
| CAGAGGGCCG CTTGCTAGCG CTCTTAATCC TAGAAATGGG GCTCTTGCCC AGTGAGATTT | 3180 |
| TAGCCATCAA GGTGCGGAC ATCAATCTGG ATTTTCAGGT GTTGCGAATC AGCAAGGCTT | 3240 |
| CCCAACAGAG GATTGTCACC ATTCCACGG CCTTGCTTTC AGAATTGGAA CCCTTGATGG | 3300 |
| GGCAGACCTA TCTTTTGAAG AGAGGAGAGA AACCCTATTC TCGTCAGTGG GCCTTTCGTC | 3360 |
| AGTTAGAATC TTTGTCAAG GAGAAAGGTT TTCCATCCTT ATCAGCTCAA GTCTTACGTG | 3420 |
| AACAGTTTAT TCTAAGACAA ATAGAAAACA AGGTCGATTT GTACGAAATT GCAAAAAAAT | 3480 |
| TAGGATTAAA AACAGTCCTG ACCTTAGAAA AATATAGATA ATGGATATTA AATTAAAAGA | 3540 |
| TTTTGAAGGA CCCCTGGACT TGCTCTTGCA TCTGGTTTCT AAGTACCAGA TGGATATCTA | 3600 |

| | | | | | | |
|------------|------------|-------------|------------|-------------|-------------|------|
| 1060 | | | | | | |
| CGATGTGCCC | ATTACGGAAG | TCATCGAACA | GTATCTAGCC | TATGTCTCAA | CCCTGCAGGC | 3660 |
| CATGCGTCTG | GAAGTGACGG | GTGAGTACAT | GGTCATGGCT | AGTCAGCTCA | TGCTGATTAA | 3720 |
| GAGTCGTAAA | CTCCTTCCGA | AGGTAGCAGA | AGTGACAGAC | TTGGGGGATG | ACCTGGAGCA | 3780 |
| GGACCTCCTC | TCTCAAATCG | AAGAATATCG | CAAGTTCAAG | CTCTTGGGTG | AGCACTTGGA | 3840 |
| AGCCAAGCAC | CAAGAACGGG | CCCAGTATTA | TTCCAAAGCG | CCGACAGAGT | TGATTTACGA | 3900 |
| AGATGCGGAG | CTTGTGCATG | ACAAGACGAC | CATTGACCTC | TTTTTGACTT | TTTCAAATAT | 3960 |
| CCTAGCCAAG | AAAAAAGAGG | AGTTTGACACA | AAATCACACG | ACGATCTTGC | GGGATGAGTA | 4020 |
| TAAGATTGAG | GACATGATGA | TTATCGTGAA | AGAGTCCTTG | ATTGGACGAG | ATCAATTGCG | 4080 |
| CTTGCAAGAT | TGTTTCAAGG | AAGCCCAGAA | TGTCCAAGAG | GTCATCACCC | TCTTTTTGGC | 4140 |
| AACCTTAGAG | TTAATCAAAA | CCCAGGAGTT | GATCCTCGTG | CAAGAGGAGA | GTTTTGGAGA | 4200 |
| TATCTATCTC | ATGGAAAAGA | AGGAAGAAAG | TCAAGTGCCT | CAAAGCTAGA | CTTGATAGAG | 4260 |
| AGGAAAGATG | AGTACTTTAG | CAAAAATAGA | AGCGCTCTTG | TTTGTAGCGG | GTGAAGATGG | 4320 |
| GATTCGGGTC | CGCCAGTTAG | CTGAACCTCT | CTCTCTGCCA | CCGACAGGCA | TCCAGCAAAG | 4380 |
| TTTAGGAAAA | TTAGCCCAGA | AGTATGAAAA | GGACCCAGAT | TCCAGTTTGG | CTTTGATTGA | 4440 |
| GACAAGTGGT | GCTTATAGAT | TGGTGACCAA | GCCTCAATTT | GCAGAGATTT | TGAAGGAATA | 4500 |
| CTCTAAGGCG | CCTATCAACC | AGAGCTTGTC | TCGGGCTGCC | CTTGAGACCT | TGTCCATTAT | 4560 |
| TGCCTACAAA | CAGCCGATTA | CGCGGATAGA | AATTGATGCC | ATCCGTGGAG | TTAAGTCGAG | 4620 |
| TGGAGCCTTG | GCAAAGTTGC | AGGCTTTTGA | CCTGATAAAG | GAAGACGGGA | AAAAGGAAGT | 4680 |
| ATFGGGGCGC | CCCAACCTCT | ATGTGACTAC | GGATTATTTT | CTAGATTACA | TGGGGATAAA | 4740 |
| CCATTTAGAA | GAATTACCAG | TGATTGATGA | GCTTGAGATT | CAAGCCCAAG | AAAGCCAATT | 4800 |
| ATTTGGTGAA | AGGATAGAAG | AAGATGAGAA | TCAATAAGTA | TATTGCCCCAC | GCAGGTGTGG | 4860 |
| CCAGTAGGAG | AAAAGCAGAA | GAGCTGATTA | AGCAAGGCTT | GGTGACGGTT | AACGGCCAAG | 4920 |
| TGGTGCGTGA | ACTAGCAACC | ACTATCAAGT | CAGGCGACAA | GGTCGAAGTT | GAAGGTCAAC | 4980 |
| CTATCTACAA | CGAAGAAAAG | GTCTACTATC | TGCTTAACAA | ACCACGCGGT | GTGATTTCCA | 5040 |
| GTGTGACAGA | TGATAAGGGT | CGCAAGACGG | TTGTCGACCT | CTTGCCCAAT | GTCAAAGAGC | 5100 |
| GTATTTACCC | TGTGGGTCGT | TTGGACTGGG | ATACATCAGG | TGTCTTGATT | TTGACCAATG | 5160 |
| ATGGGGACTT | TACAGACGAG | ATGATTCACC | CTCGTAATGA | GATTGACAAG | GTTTATGTCTG | 5220 |
| CGCGTGTTAA | AGGTGTGGCC | AATAAGGACA | ATCTCCGCCC | CTTGACCCGT | GGTCTTGAGA | 5280 |
| TTGATGGTAA | GAAAACCAAG | CCAGCTGTTT | ATGAAATTCT | CAAAGTGGAC | CCAGTCAAAA | 5340 |
| ATCGCTCTGT | GGTGCAGTTG | ACCATCCATG | AAGGGCGTAA | CCATCAGGTT | AAAAAGATGT | 5400 |

1061

| | |
|--|------|
| TTGAAGCTGT TGGTCTCCAA GTAGATAAGT TGTCTCGGAC TCGTTTCGGA CACCTAGACT | 5460 |
| TGACAGGACT CCGTCCAGGA GAATCCCGTC GTCTTAATAA AAAAGAAATC AGCCAACTAC | 5520 |
| ACACCATGGC TGTAATAAG AAATAATGAA ACGAATTTTA ATAGCGCCTG TCGCGTTTTA | 5580 |
| CCAACGTTTT ATCTCACCAG TCTTTCCACC CTCTTGTCGC TTTGAGCTGA CTTGTTCCAA | 5640 |
| CTACATGATT CAGGCTATTG AAAACATGG GTTTAAGGGG GTATTGATGG GCTTGGCTCG | 5700 |
| GATTTTACGT TGTTCATCCCT GGTGAAAAC AGGTAAGGAC CCCGTTCCAG ACCGCTTTTC | 5760 |
| CCTTAAACGA AATCAAGAAG GGGAAATGAGG TGGGGTAAAT AGATTTCAAA ATGATAAAAA | 5820 |
| CGCATCCTAT CAGGTTTGAG TGAACCTGAT AGGATGCGTT TTAGAATGTC AAAATTTTAT | 5880 |
| ACTCTTCGAA AATCTCTTCA AACCGCGTCA GCTTTCATCT GCAACCTCAA AACAGTGTTT | 5940 |
| TGAGCAACCT GCGGCTAGTT TCCTAGTTTG CTCTTTGATT TTCATTGAGT ATTAAATTGA | 6000 |
| GTTTGAAGTG GCTTATTTCA AAGCTTTTTG TATGCTTCA ATCATGAGTT TTGTGTATTC | 6060 |
| AAGTCCGCCT CCGCTTAGAT ACCAGAGGTC TGGTGTTAGT TGGATAATCT TACCATTTT | 6120 |
| AGCAGCAGGT GTTTCAGCGA TAAGGGCATT TTCTAGGACA CCGTCGTTGC TAGAGTTGTC | 6180 |
| CCCACCGATG GCAAGGGTAC GGTGATGAC AAAGAGGATG TCAGGGTTGA TTTCTTTGAC | 6240 |
| ACTTTCAAAG CTGACTTCTT GTCCGTGGCG TGAGTCTTCA AATTTTGATC CAGTTGGTTT | 6300 |
| GAATTTCAAG GTTTGGTACA AGAAAGAGAA ACGAGATTG GCACCAAAGG CTGCCATTTT | 6360 |
| TCCTTCATTA AGGAGGATCG CAAGGGCTTT TTTGTCAGAG CTTTCATTTT TAGTAGCGAC | 6420 |
| TTCTTGGATG CTCTTGCTCA GCTTGGTCAA TTCTTCCTTG GCTTTCTGTG TACCAGTTTC | 6480 |
| GCCGAAGGCA CTTGCTAAGG ATTCGATATT AGCCTTGGA GAAGTCCAGT AGTCGTCCTT | 6540 |
| GCTTGCTTGG AAGAGAACGG TTGGGGCGAT TTCTTTGAAT TTGTCTACGA ATTTTGTGT | 6600 |
| ACGTGGCGAA GCGATAATCA AATCAGGCTC AAGGGCGGCG ATAGCTTCTA AATCAGGTTC | 6660 |
| TTTCATAGAA CCAACATTTT TGACAGTTCC CACTAGGTCT TTTAGATAAG TCGGAACAGT | 6720 |
| TTTTGTAGGC ATTCGACGA TATTTTTTTC AAATCCTAAA GCGCGAATAG TATCCGCAGC | 6780 |
| GCCGAGGTCA AAGGTCACAA TCTTTTCAGG AACTTTGGAA AGTTTGACCT CGTCCAGTGA | 6840 |
| ACTTTTAATG GTTACCTCTG TTGGAGCAGA GCTACTGGTC TCTGTCTGAC TAGTGCTTGA | 6900 |
| GTTTGTACTA CATGCACCAA GTAGGAGCAA GAAGCTGGCC ACTAGGGCAG TGAAATAAAG | 6960 |
| TTTAAGGGAT GTTTTCATAA TTTCTCCTTT TTAATAATGTG ATAACGATTT AGGGAGTCTC | 7020 |
| TTAATCTTAT TGACTAAGAG ACTGAAGGTT CTCTAACTTG AGCTTTTATG TTACTAGCTA | 7080 |
| TAGATACAGA TCTTTTGTG ATTGATATCA GCTAGCGTGA TGGGAATCTC ATAAAGTTGA | 7140 |

1062

| | |
|---|------|
| CTGAGCAGGT CAGCCTGCAT GATTTGATCG GTTCTTCCCT TGCTAAAGAC CTGGCCGTCC | 7200 |
| TTGAAGGCCA CAATTTCATC TGCATACTGA CTGGCCATGT TGATATCGTG GAGGACGATG | 7260 |
| ATAATGGTCT TGCCGAGTTC CTCCACCACT CGTCGAAGAA TCTGCATCAT GCTGACGCTT | 7320 |
| TGCTTGATAT CGAGATTGTT GAGTGGTTCG TCCAGCAAGA TAAAGTCCGT ATCCTGGGCC | 7380 |
| AGTACCATAG CGATAAAGAC GCGCTGGAGT TGCCCCCTG ACAGGCTATT GATGTAGCGG | 7440 |
| TCTTTTAAGT TGGTCAGTTC TAAATAGTTC AGAGTTTCTC GGATTTTTC CCAGTCTTCT | 7500 |
| GATCTAAGTC GACCTCGGCT GTAGGGAAAA CGTCCAAAAC TGACCAGTTC TTCAACAGTC | 7560 |
| AATTTGGCTT GGTAATTGAT TTTCTGTTTT AGGATGGTTA GTTCTTGGGC CAGTTCTTGC | 7620 |
| GAATTCCAGC TCTCGATTTC ACGTCCTTTG ATACTGAGAA CTCCCTGATC TTTCTTGGTT | 7680 |
| AGCCTGCTCA TGATGGAGAG GAGAGTCGAT TTTCCAGCAC CATTTGGACC AATAAAGGCT | 7740 |
| GTCAGTTTTT GAGGACTGAC TTCAAGCGAA ATGCCCTGCA AAATATCCTG TTTTGAATG | 7800 |
| GATTTGTCAA TGTTTCCAG TTCACTGAC GAGACCTCCT ATATAGTAAG ATAAAGAATA | 7860 |
| AGAAGCCACC CACACTCTCA ATGATCATAC TGATACGAAT TTCCAGTGCA AAGACTCGTT | 7920 |
| CAATCAAGGC TTGCCCCAAG GTTAAGCTAA TAAATCCAAC CAGAATGGCC ACTATAAAGA | 7980 |
| GTAACCTGTG CTGATAGTCT TTGACAATCA GGTAGGTGAG GTTGGCCAGT ATAAAGCCGA | 8040 |
| AGAAGGCCAT AGGTCCTACC AAGGCAGTGG CCGTTGAGGT CAAAAGCACG ATTCCCCAGA | 8100 |
| GGAGCTCTTT CTGTTCTTTT TCAACATCGA GTCCCAATAT CTGAGCCGTT TCTCTTTGCA | 8160 |
| GGTGCAAGAC ATCTAGAACG ACTGCTTTTC GAAAGAAAAA GATTGTCAA CCGAGGATGA | 8220 |
| TCAGAGAACC GATGGCTAGG ATGGAAGTGT TGAGATGTTG AAAGGAGGCA AAAAGACTAT | 8280 |
| TTTGCAGTTT ATCGTATTCG TTTGGATCCA TTAGGACTTG AAGGAAGGTG CTGATATTTC | 8340 |
| GAAAGAGACT TCTGAGCGCT AGACAGATCA GCAGGACGAA GACCAGGTCT TGCTTCATCA | 8400 |
| GTGTCTTCAA GTAACCTTGT AAGGCGAGAA AGAAGAGGGA CTGGACAAGA AGTAAGACTA | 8460 |
| GGAATTCTAA GATAGGGGAT TTGCCAAGTT GAAGAACTT GCTTTCAAAA ACCAGTAGTA | 8520 |
| GGGTTTGTAG TAGGACGTAG AAGGATTCAA TTCCCAAAAT ACTAGGCGTC AGGAAGCGAT | 8580 |
| TTTCCGTCAG GGTTTGAAAA CTAATGGTCG AAATCCCACT CGCGATGGCT ACCAAGAGAT | 8640 |
| AAACGATGAT CTTTGGGAA CGCAACTTCC AAGCAAAGGC TGACAAGTGA GTGATGGGCC | 8700 |
| AAAAGTAGAG AAGACAAGCT CCGATGGCAA GAATAATGAG AATCCAGAAG AGCTTGGTAT | 8760 |
| GTTTGCTTTT AGTCTGCATC TTTTCGTCCC CCTCTCCAGA GAAGTAGGAT AAAGACGAGA | 8820 |
| CTACCGATGA TTCCTAGCAA GAGACTGACA GACAACTCAT AGGGCCTAAT CAGAACTCGG | 8880 |
| GATAGGATAT CGCAAGCCAG AACTAGATTG GCACCAACCA GTGCGACCAT GAGTTTGGTT | 8940 |

1063

| | |
|--|------|
| TGACTTAGAT TATCTCCATA GCGCTTGCGA ACAAGATTGG GAACGATAAC TCCGAGAAAT | 9000 |
| GGTAGGCCAC CCACGGTAAT CATGGTGACG CTTGTCGTTA GCGCCACCAG AAAGAGGGCC | 9060 |
| AGTTTTTCAA GTAGGGAGTA GGAAATCCCC AAACCTCTCGC TGGTTTCTTT CCCTAGATTC | 9120 |
| ATGATGGTGA AGGTTTGGA TAATTTCCAA ACGGTTATCA GGATGATGAG GCCTAAGAAG | 9180 |
| AGCCACTCAT ACTGATGGGT CTGAATCATG GAGAAGGAGC CCTGGGTCCA GGCAGTCATA | 9240 |
| CTCTGAACCA GATTGAAACG ATAGGCGATA ACTTCTGTGA CTGAGCCGAT AATCCCGCTA | 9300 |
| TAGATGATCC CAATCAGAGG CAACATCCAC CTTTCCTTTA CAGTAAAAAT GGTCATAAAG | 9360 |
| GCTAGGAAGA AGAGGGTGAA TACGATGGAT GAAACAAAAG CGAAGAGCAT CTTGTGGGTC | 9420 |
| AGACTAGCCG ATGGAAAGAC AAAAAGGCTC AGCACCATTC CCAGTTTGGC GGCTTCAGTC | 9480 |
| GTTCCAAC TG TACTCGGTGC AGCAAACTGA TTTTGGGTAA TAGTCTGCAT GAGAAGGCCT | 9540 |
| GCCATACTCA TACTAGAGGC AGTCAGGAGA ATACTGATAG TTCTTGGGAG ACGGGACTCT | 9600 |
| TGAAAGAGGA GCCAGGTCTG CTGGTCGAAA TCAATAGCT TTCCCATGA AAAATCACTG | 9660 |
| GTCCCAATGC TAATAGAGAG AAAGACTAGG AGTAGAAGTA AGCCAGG | 9707 |

(2) INFORMATION FOR SEQ ID NO: 165:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 5910 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 165:

| | |
|--|-----|
| CCGCAATTAT GCTTGAAAAG GAGTATACTT ATAAGTAACG CAAACGTTTG CGTCTGAAAA | 60 |
| ATACGCAACG TTCCATTATT TTAACACACG AGGTGCTATT ATGAAAAAAC GTCAAAGTGG | 120 |
| TGTGTTGATG CACATCTCTT CTCTCCAGG AGCTTACGGA ATCGGATCAT TTGGTCAAAG | 180 |
| TGCTTACGAC TTCGTTGATT TCTTGGTCCG TACAAAACAA CGTTACTGGC AAATCCTTCC | 240 |
| ATTAGGAGCA ACTAGTTACG GGGATTCTCC TTACCAATCT TTCTCAGCCT TCGCAGGAAA | 300 |
| CACTCATTTT ATCGATTAG ATATCTTGGT GGAGCAAGGT TTGTTGGAAG CAAGTGACCT | 360 |
| TGAAGGAGTT GACTTTGGTA GCGATGCGTC TGAAGTTGAC TATGCTAAAA TCTACTATGC | 420 |
| ACGTCGTCCT CTTTGTAGAAA AAGCGGTGAA ACGTTTCTTT GAAGTCGGAG ATGTTAAAGA | 480 |
| TTTTGAGAAA TTTGCTCAAG ACAACCAATC ATGGCTTGAG CTCCTTGCTG AGTATATGGC | 540 |
| TATCAAAGAG TATTTTGACA ATCTTGCTTG GACTGAATGG CCAGATGCAG ATGCTCGTGC | 600 |

1064

| | |
|--|------|
| TCGTAAAGCT TCAGCACTTG AAAGCTATCG TGAGCAATTG GCAGACAAGT TGGTTTACCA | 660 |
| CCGTGTGACT CAATACTTCT TCTTCCAACA ATGGTTGAAA TTGAAAGCTT ACGCTAACGA | 720 |
| CAACCACATC GAAATCGTTG GGGACATGCC AATCTACGTA GCGGAAGATT CAAGTGATAT | 780 |
| GTGGGCAAAT CCACATCTCT TCAAAACAGA TGTCAATGGT AAGGCTACTT GTATCGCAGG | 840 |
| ATGCCCAACA GATGAGTTTT CTGTAACTGG TCAGCTTTGG GGTAAATCCAA TCTATGACTG | 900 |
| GGAAGCAATG GACAAAGACG GCTACAAATG GTGGATTGAA CGCTTGCGTG AAAGCTTCAA | 960 |
| AATCTACGAT ATCGTTCGTA TCGACCACTT CCGTGGCTTC GAATCTTACT GGGAAATCCC | 1020 |
| TGCTGGTTCC GATACAGCAG CACCTGGTGA GTGGGTGAAA GGTCCAGGTT ACAAGCTTTT | 1080 |
| TGCAGCCGTT AAGGAAGAAC TTGGTGAGCT AAACATCATC GCAGAAGACC TTGGCTTCAT | 1140 |
| GACAGATGAA GTGATCGAAT TGCGTGAACG TACTGGCTTC CCAGGAATGA AGATTCTTCA | 1200 |
| ATTTGCCTTC AACCCAGAAG ACGAAAGCAT TGATAGCCCA CACTTGGCAC CTGCTAACTC | 1260 |
| AGTTATGTAC ACAGGAACAC ACGATAACAA TACGGTTCTT GGTGGTACC GTAATGAGAT | 1320 |
| TGATGATGCG ACTCGTGAGT ACATGGCTCG TTACACGAAC CGTAAAGAAT ACGAAACAGT | 1380 |
| GGTACACGCT ATGCTTCGTA CAGTATTTTC ATCAGTTAGC TTTATGGCAA TTGCAACTAT | 1440 |
| GCAAGATTTA CTAGAATTGG ATGAGGCAGC TCGTATGAAC TTCCCATCTA CCCTTGGTGG | 1500 |
| AAACTGGTCT TGGCGTATGA CTGAAGATCA ATTGACACCA GCTGTCGAGG AAGGTTTGCT | 1560 |
| TGACTTGACA ACAATTTATC GCCGAATTAA TGAAAATTTG GTAGATTTAA AGAAATAAGA | 1620 |
| CAATAATCAG GAGACAACATA AACATGTTAT CACTACAAGA ATTTGTACAA AATCGTTACA | 1680 |
| ATAAAACCAT TGCAGAATGT AGCAATGAAG AGCTTTACCT TGCTCTTCTT AACTACAGCA | 1740 |
| AGCTTGCAAG CAGCCAAAAA CCAGTCAACA CTGGTAAGAA AAAAGTTTAC TACATCTCAG | 1800 |
| CTGAGTTCTT GATTGGTAAA CTCTTGTCOA ACAACTTGAT TAACCTTGGT CTTTACGACG | 1860 |
| ATGTTAAAAA AGAACTTGCA GCTGCAGGTA AAGACTTGAT CGAAGTTGAA GAAGTTGAAT | 1920 |
| TGGAACCATC TCTTGGTAAT GGTGGTTTGG GACGTTTGGC TGCCTGCTTT ATCGACTCAA | 1980 |
| TTGCTACTCT TGGTTTGAAT GGTGACGGTG TTGGTCTTAA CTACCACTTT GGTCTTTTCC | 2040 |
| AACAAGTTCT TAAAAACAAC CAACAAGAAA CAATCCAAA TGCATGGTTG ACAGAGCAAA | 2100 |
| ACTGGTTGGT TCGCTCAAGC CGTAGCTACC AAGTACCATT TGCAGACTTT ACTTTGACAT | 2160 |
| CAACTCTTTA CGATATTGAT GTTACTGGTT ATGAAACAGC GACTAAAAAC CGCTTGCGTT | 2220 |
| TGTTTGACTT GGATTCAGTT GATTCTTCTA TTATTAAAGA TGGTATCAAC TTTGACAAGA | 2280 |
| CAGATATCGC TCGCAACTTA ACTCTCTTCC TTTACCCAGA TGATAGTGAC CGTCAAGGTG | 2340 |
| AATGCTCCG TATCTTCCAA CAATACTTCA TGGTTTCAAA CGGTGCGCAA TTGATCATCG | 2400 |

1065

| | |
|--|------|
| ACGAAGCAAT CGAAAAAGGA AGCAACTTGC ATGACCTTGC TGA CTACGCA GTTGTCCAAA | 2460 |
| TCAACGATAC TCACCCATCA ATGGTGATTC CTGAATTGAT TCGTCTTTTG ACTGCACGTG | 2520 |
| GTATCGATCT TGACGAAGCA ATCTCAATTG TTCGTAGCAT GACTGCCTAC ACTAACCACA | 2580 |
| CAATCCTTGC TGAAGCGCTT GAAAAATGGC CTCTTGAATT CTTGCAAGAA GTGGTTCCTC | 2640 |
| ACTTGGTACC AATCATCGAA GAATTGGACC GTCGTGTGAA GGCAGAGTAC AAAGATCCAG | 2700 |
| CTGTTCAAAT CATCGATGAG AGCGGACGTG TTCACATGGC TCACATGGAT ATCCACTACG | 2760 |
| GATACAGTGT TAACGGGGTT GCAGCACTCC ATACTGAAAT CTTGAAAAAT TCTGAGTTGA | 2820 |
| AAGCCTTCTA CGACCTTTAC CCAGAAAAGT TCAACAACAA AACAAACGGT ATCACTTTCC | 2880 |
| GTCTTGGCT TATGCATGCT AACCCAAGAT TGTCTCACTA CTGGATGAG ATTCTTGGAG | 2940 |
| ATGGTTGGCA CCATGAAGCA GATGAGCTTG AAAAATTTT GTCTTATGAA GACAAAGCAG | 3000 |
| TTGTCAAAGA AAAATGGAA AGCATCAAGG CTCACAACAA ACGTAAATTG GCTCGTCACT | 3060 |
| TGAAAGAACA CCAAGGTGTG GAAATCAATC CAAATCTAT CTTTGATATC CAAATCAAAC | 3120 |
| GTCTTCACGA GTACAAACGC CAACAAATGA ACGCTTTGTA CGTGATCCAC AAATACCTTG | 3180 |
| ACATCAAAGC TGGTAACATC CCTGCTCGTC CAATCACAAT CTCTTTGGT GGTAAAGCAG | 3240 |
| CTCCAGCCTA CACAATCGCT CAAGACATTA TCCATTTAAT CCTTTCATG TCAGAAGTTA | 3300 |
| TTGCTAACGA TCCAGCAGTA GCTCCACACT TGCAAGTAGT TATGGTTGAA AACTACAACG | 3360 |
| TTACTGCAGC AAGTTTCCTT ATCCCAGCAT GTGATATCTC AGAACAAATC TCACTTGCTT | 3420 |
| CTAAAGAAGC TTCAGGTACT GGTAAATGA AATTCATGTT GAACGGAGCT TTGACACTTG | 3480 |
| GTACTATGGA CGGTGCTAAC GTGAAAATCG CTGAGTTGGT TGGAGAAGAA AACATCTACA | 3540 |
| TCTTCGGTGA AGATTCAGAA ACTGTTATCG ACCTTTACGC AAAAGCAGCT TACAAATCAA | 3600 |
| GCGAATTCTA CGCTCGTGAA GCTATCAAAC CATTGGTTGA CTTTCATCGTT AGTGATGCAG | 3660 |
| TTCTTGCAGC TGGAAACAAA GAGCGCTTGG AACGTTTTTA CAATGAATTG ATCAACAAAG | 3720 |
| ACTGGTTCAT GACTCTTCTT GATTTGGAAG ACTACATCAA AGTCAAAGAG CAAATGCTTG | 3780 |
| CTGACTACGA AGACCGTGAC GCATGGTTGG ATAAAGTCAT CGTTAACATT TCTAAAGCAG | 3840 |
| GATTCTTCTC ATCTGACCGT ACAATCGCTC AGTATAACGA AGACATCTGG CACTTGAAC | 3900 |
| AATACTCTTC GAAAATCTCT TCAAACCACG TCAGCTTTAT CTGCAACCTC AAAGCAGTGC | 3960 |
| TTTGAGCAAC TGC GGCTAGC TTCCTAGTTT GCTCTTTGAT TTTTATTGAG TATAAGATAC | 4020 |
| AAATTTATAC TAATACATTT TGTAAAAAG CGAGTTTCGA TTGAAATTCG CTTTTTTAAT | 4080 |
| GATGTAGATT TGGGTCAATC TTGTCTAAAA ATAGGGAAAT CCTAGATACA GTGAAGGCTT | 4140 |

1066

| | |
|---|------|
| TAAATGCTGG TTTTACTGT CCTCAGCCTT ATATTTTTTC GTAGTTGGTT ACCTCATATC | 4200 |
| TATTATATTC GCTTACATAA AGTATTATAA TATAATTGTA GGAAAGAAGG TGTMTTATG | 4260 |
| ATATACACAC TTAAATGGT GTTGTATTAT ACCTTCTTG TAATAAGCTT GTTACCTGAT | 4320 |
| AAGATTTTGG GAAAAATAA AAAAATTGG AAAATAGTTT TTGCAATATT GACGGCAGTG | 4380 |
| GCAGCATTGT CATTTATGTA CTAAGTTATT TTAAGAATGT AGGGAAATAA ACCCTACATT | 4440 |
| CTTTTATGTT TTTTCTGTTT TCTAAATCTT ATTTATCCAA GCGATTCAAC ATTTCTTGCT | 4500 |
| TCTTCGCTTC AAGTTCTGCA CGCTTTTCTT CGATTCGGC ATGTTTTTTC TCGAGTTCAG | 4560 |
| AACAACCTGC ACCATTGCTA AATTCCTTTC GCCATCAGGA GATAGGGTGA GTCGACATGT | 4620 |
| CTATTACTCA CCCAAAGCAG TCCTACAAAG CAGGAATTTT CTGTTACTTT TTTGGAAATA | 4680 |
| GTAACGTTTA TACAGCTTTG AACTTCGTA TCAAAGCGCC AAACACACTC CGAGGGGTTT | 4740 |
| ACAGAAAGCA GAAAAGGAAT GATCTGGTAT AAGATCATTC CTTTTCyCTC TTTTCTTTA | 4800 |
| AGTAATTATA TACAATGTAC GACGAAGTCG TCATTGCAAT GCTGATCCAC CACCTAAAGG | 4860 |
| GAACCTTAAA CAACATTGAT AAGATAAAGA ATATAAACAA CGAAAATACG TTATACCCAA | 4920 |
| TTAATTTTAT TGTATATCTC ATGATTAAAA GTTAATCCTT CCGTTGTTAG GAATGGCATC | 4980 |
| ATTTTTATCC CATAATTGTG CTAAATAAGT CCCCCTGAT AATAAATTCA TAGCGAATTC | 5040 |
| TAAAGCAACA TCATTTACAA ACCAACTACC TAGATATCTA GAAATGCTG AACGAATAGC | 5100 |
| ACTTTTTGCT GCATGTTTTC CTTTTACTTT AATTAGATTT GCAAGGCCTG CAGTAGTTCC | 5160 |
| TCCTAATGCT AAAGCTATTG CAGTATCTAA TAGAGCACCC ATTTGATTAA CTGTAATACC | 5220 |
| TTGCCAAACT GCTCTAAATG GAGAGTATGT AGGTGGGATT GTATAATCGC CTTGTAATTG | 5280 |
| TCGGTTAATT ACTTCTTTGA TCCATTGTTG TGAGACGTCT GGATGAAAAG ATTGGATTTC | 5340 |
| GTTTGCAAGT GTATTGATTT GTTCTTCTGT TAGAGAAGTG ACAGGTTGAA GTTCCATATT | 5400 |
| TGTTTCAATT TGTGATACTT GTTCAGAAGC GTATACAGCT GAAACACTTG GAATCGCTGA | 5460 |
| TACAATTAAC ACAATTGACG TCAAAAAAAC CGAAATAAAT TTCATTAAAT TGTTCATGAG | 5520 |
| CTTTTCTCCT TTTTATTTGC ATCTGCTTAC ATTTTATCAT ATACTGTTAT TATAGTCAAA | 5580 |
| AAAATATGCT ATTATGTTAA AAAAATATTT TTCAAAATAT AAATGGACGG ATTTATTTTG | 5640 |
| GATTTTATTT GTTATTTTGA CCTGCCTCTA TATTGGTAAC CATGATTTGT TTACTCTCAA | 5700 |
| TCATCAAGAA TTCTCTTTTC GTGGTAGCGT TTGGGGTCTG GTACTGGCCT TATATCACTT | 5760 |
| ACTATTCATT GATAAGTTTG TTATATCGAA TCGAAAATAA AGATTAGAGC TATGCTTGAC | 5820 |
| TGTGTACTTT TAGGATTTAT TTTGGAGGAA GATTTTGTCT CTATTATTTA TTATTTTAAA | 5880 |
| TTTATTTATT TTGTATAAGA TCTATTCTTT | 5910 |

1067

(2) INFORMATION FOR SEQ ID NO: 166:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5406 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 166:

```
GGCATAGCGA CTCATTTTTT CAACTGTCCA GGCTGGATAC CAGACTAATT TAACCTCAGT      60
ATCCGTTACT TCTGGAACCT CTATCATAGC ATCATAAATC TGGTCTGTCA AAAGGTCTGC      120
TAAGGGACAA CCCATAGTTG TCAAAGTCAT GTCAATCTCT GTTTGCCCTG TGTACCCGTC      180
AAAACGAATC TCATAGATCA AACCAAGATT GACAATATCG ATTCCCAACT CAGGGTCGAT      240
GACTTCTTCC AAGGCTGTTA AAATCCGTGT TTTGATGTTT TCAATTGCTT CTTCTGTATA      300
AGCCATATTT TCCTCACTCT TAGTCTTCAA TAAATCACG AAGCGGTTTG CTACGACTTG      360
GTTGGCGTAG TTTTCTCAAA GCCTTTGCTT CAATCTGACG GATACGCTCA CGAGTTACGT      420
TAAAGACTTT CCCACATCT TCAAGTGTGC GCATTTTCC ATCATCTAGT CCAAAACGTA      480
GACGCAGAAC ATTTTCTTCA CGGTCTGTAA GAGTATCTAA GATTTTCATCC AATTGCTCAC      540
GCAAGACGAT ACGAGTCGTA TAATCCACTG GATTTTCAAT CACTTCATCT TCGATAAAGT      600
CTCCAAGGTG GCTATCGTCC TCTTCACCGA TAGGAGTTTC AAGAGATACT GGTCTTGGG      660
CAATCTTCAA GATTTACGTA ACCTTATCAG GTGTCATATC CATTGCTTCA GCAATCTGTT      720
CTGGTGTCGG ATCTTGCCCC AATTCTTGAA GGAGATTCCG CTGTTACGTA ACCAATTTAT      780
TGATAGTTTC AACCATGTGA ACTGGGATAC GGATGGTACG AGCTTGGTCC GCAATAGCAC      840
GAGTGATAGC CTGACGAATC CACCAAGTTG CATAAGTTGA AAACCTGAAC CCTTTAGAAT      900
AGTCAAACTT GTCAACCGCC TTCATCAAGC CCATATTTCC TTCTTGAATC AAGTCAAGGA      960
ACTGCATACC ACGACCGACA TAGCGTTTGG CAATGGAAAC AACCAAACGA AGATTGGCTT     1020
CCGCAAGACG TTGTTTGGCT TCGATATCAC CAGCTTCAAC AGCCAGTGCC AACTCTTTCT     1080
CCTCTTCATT GGTCAAGAGA GGAACGACCC CTATTTCTTT CAAGTACATA CGGACAGGGT     1140
CATTGACCTT AGCAGAAGTT GACCCAATCA AGTCCTCATC GCTGAGTTCT GGTCTTCTTT     1200
CATTGCTGAG AACACGCGCA CTTGGATTTC CTTGTTATC TGTGATAGAA ATGCCTGCAT     1260
CCTGAATCCG TTGCAAGAGA TCTTCAATCC CATCAGCGTC CAAGGTAAAA GGAATAACCA     1320
GACTTGCATT GATTTTCATCA TCTGTGCTG TCCCTTTTGT CTTATGATTA CGGATAAATT     1380
```

1068

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| CTGCTACCTG | TACGTCAAAT | GTTGTTACTT | CTTTTGTGTT | TGTTGCCATT | ATTACTCCAT | 1440 |
| TCTTCTCTTT | TGGGAAATTA | AACGTTCCAA | TCTTCTAGG | GCTGTATCTG | TATCTCCTAC | 1500 |
| ATGGCTAGCT | TCCTGCACCT | TCTTTTGTAT | TCTCATATTG | TCCTGATTCA | AGAGAGCCTT | 1560 |
| GTTTCGAGTC | ATCTCTACTT | CACTAAGTTC | CTGCGGCGAT | ATCTCAGCAG | GCAAATCCTG | 1620 |
| AGCTAAAACT | TGGTACCAAG | CTCTTTCAAC | TTCTCTGTG | TGCTCTGCTA | AAACTTCTGG | 1680 |
| AGGAAGATTT | CCATACTGGC | CAAGCAAGTC | ATATAAGACC | TGAAATTCAG | GTGTAGCAAA | 1740 |
| TGCAAAGTCT | TCTCGCAAAC | GGTAATCGTT | CAAAACAAGA | GGGGATTCCA | TCATCCGATA | 1800 |
| GAGTAGATGG | GCTTCTGCCC | TCATAATAGC | CGATAACTGC | TTGGTGACAG | GCATGGTGAT | 1860 |
| TGGCGTCGGT | CTGGAAATTC | CTTCCATGCG | ATTCTGCCTT | TGCACCTGAC | GACTCTCATT | 1920 |
| AACAATCTGC | TCAATCTGGG | TATAATCAAA | GGACGCCAGA | CTGTCAGCTA | AAATATGAAT | 1980 |
| ATAGCTGTTT | TGAGCAGCGA | TGGACTTTTC | TTGAACAATC | AAGGGAGCTA | TTTTTTCAAG | 2040 |
| AAACTCAATC | TGAGCCTGCA | GATTTTCACT | GTTTTCAGGT | TTGTACTGAT | GAATGTAGAA | 2100 |
| CTCAATCGGA | CTAATACGAG | TTTTCGTTAA | TAGATAGGCC | AAGTCTTCTG | GACCATTTTT | 2160 |
| TTGTAGATAC | TCATCAGGAT | CCAAGTTATC | AGGCATGCTG | ACGATTTGCA | CAGGCATATC | 2220 |
| ACCAATTTCA | TCCAATGCTT | TCAATGTCGC | GGCTTGCCCA | GCCTTATCTC | CATCGTAAAC | 2280 |
| AAGAACCAAT | TTCTTGTTTA | ACCTTTTCAG | ATGCTCAACA | TGCTCTCGAC | TCAAGGCTGT | 2340 |
| TCCCATCGAC | GCCACAGCAT | TTTCGATTCC | AGCCCGATAG | GCTGCAATAA | CATCCATGAA | 2400 |
| TCCTTCCATC | AGGTAAATCT | CACTAGCTTT | TCCAGAAGAT | CTTTTGTCCC | TATCCATATG | 2460 |
| ATATAATTCG | TAACTTTTGT | TAAAAATTGC | AGTCGATCGG | CTGTTTTTAT | ACTTAGAAGT | 2520 |
| TTGTGAATCC | GTTTTTTGCC | AGATACGACC | TGAGAAGGCA | ATGACCTTTC | CTTGGTCATT | 2580 |
| TGTCAGGGGA | AACATAATGC | GATTGTGAAA | GGTGTCTACA | AATTGATTGG | CATCCGAGAG | 2640 |
| ATAAACACAG | CCTGAATCCA | GTAATCCCTC | TTACAGATAC | TGATCAGACA | AACGTTGATA | 2700 |
| GAGATAGTTT | CGTTCTGGAG | GTGCTAAACC | AATCCAAAAA | TGTTTAAGCA | CTTCATCTGT | 2760 |
| CAACCCCGGC | TGATAAAGGT | AATTTCTGGC | CTCTTCGCCC | ATAGTCGTTG | TCATGAGAAT | 2820 |
| AGCATGGTAA | AATTTGGCTG | CATCTTCGTG | CATATCATAA | AGAGCTTGGT | GAGGTGAGGC | 2880 |
| TGACTTCTGC | TCACTATAAA | GCGGTTTTTC | AACCTCAATT | CCAACACGCT | GACCTAAGAT | 2940 |
| TTGGACTGCT | TCTATAAAGG | GAACCCCTTG | GTAATCCTCG | ATGAACTTAA | AGACATCACC | 3000 |
| TGAGCGACCA | CAACCGAAAC | AGTGATAAAA | CTGCTTGTCC | TCTACAACAT | TGAAAGATGG | 3060 |
| TGTTTTTTCA | CCATGAAAAG | GACAGAGCCC | TAGATAGTTC | CGTCCTGCCT | TTTGTAAGA | 3120 |
| AATCACATCT | CCTATGACTT | CCACAATGTT | GGCATTGTTT | TTGATTTCTT | CAATGACTTG | 3180 |

1069

| | |
|--|------|
| TTTGTCAACC ATACACAATA CCTCCATGTT ATCATAGTTT ACTTTATATA GTATACTTTA | 3240 |
| TTTCAGAAAA AAAGTAAACC ATTTCACTCA TTTTCCCTAC TTTATTCAAA GAGTTGATAA | 3300 |
| TAATCAGAGA TTTTCATTTT TGCTTTTCT TCTTGTTTA AATCTTGGAT AATTCGTCCT | 3360 |
| TCTTTTCATGA CAATCAAGCG ATTGCCGTAT TTGAGAGCAT CTTCCATATG ATGAGTAATC | 3420 |
| ATAAGGGCTG TTAGCTGATC TTTCTTAACA AATTCATCTG TCAATTCCAT CAAAGCAACA | 3480 |
| CTAGTCTTTG GATCCAGGGC AGCAGTATGC TCATCTAACA GGAGTAATTC AGGTCGCTTC | 3540 |
| AAGGTTGCCA TCAAGAGACT CAAAGCCTGT CTTTGTCCAC CTGATAAGAA CTCAATCGGT | 3600 |
| GTATTCAAGT GTTCTCAAG ACCATTTCTT ACTTTTTCAA TGGTTGCCTG AAATTCATCC | 3660 |
| TTATAGCTAG TCAAGCGTCG TGGTAACAAT CCACGCTTTT CACCACGAAA CTTGGCGATT | 3720 |
| AAAAGATTTT CAGCGACCGT CATACGGGGA GCTGTCCCCA TCTTTGGATC TTGGAAGACA | 3780 |
| CGAGACAGGT ACTTGGCAGC CTTCTCGGGT GAAACTTAG TGAGATCTTC ACCTAAAATA | 3840 |
| CGGATAGTTC CACTAGTTAG TGATAAGGTC CCTGCTATAG TGTAAAGAG AGTTGATTTT | 3900 |
| CCAGCACCAT TTCCGCCCAA AATCGTGATA AAGTCCCGTT CAAAAATTTC TAAGGAAACA | 3960 |
| TCATTTAAAA TAATCTTTTC TTCATCAAAG CCATTTTAA CGATTTTGGT TGCATTTTTT | 4020 |
| AATTCTACAA TTGCTGTCAT TTGCTTAACT TGGCTCCTTT CAAGATTGTT TGCTTAAATG | 4080 |
| TTGGAATCAT GAGGCAGACT GTAAAATCA AGGCACTGTA TAAACGAAGG TAACCTGTAT | 4140 |
| TAAAGCCAAG TGCGATAACT GCCCACACTA AAAATTGATA AGCGATAGAA CCTACAACGA | 4200 |
| TAGTAACCAA ACGCTCTGCC AAGCTCAAAC TCTTGAAAAT AACTTCTCCA ATAATCAAAC | 4260 |
| TTGCAAGCCC CACAACGATA ACCCCGATCC CTCGAGACAC ATCGGCATAA CCTTCTTGCT | 4320 |
| GAGCAATGAG GGCACCTGCA AGGGCAATCA CACCATTGTA TAAGACCAAG CCCATGAGCT | 4380 |
| CCATGCGTCC AGTATGAATC CCGAACTTC TAGCCATATC AGGATTATCC CCTGTAGCAA | 4440 |
| TATAGGCTTG TCCGAGTTTA GTGTCCAAGA AAAAGAGCAT GAGAGCAATA ACAATACTCA | 4500 |
| CAAAGATGAG ACCTGTCAAG AGTTGATTCA AATCCGAATC AAAAGGCAAA ACATCCTGAA | 4560 |
| TTTGCTTGGT TCCAAGCAGG CCTAAATTCG CACGTCCCAT AATCAAGAGC ATGATTGAGT | 4620 |
| GACAAGAAGT CATCACCAA ATCCCTGAGA GCAAGGTTGG GATCTTCCCT TTTGTATAAA | 4680 |
| GAAGGCCTGC TGCCATTCCA GCCAAACAAC CTGCTCCTAC AGCAACAAGT GTCGCTAAAA | 4740 |
| ATGGGTTCAC GCCTTTGGTT ATCAAAGTGA CAGCAACAGC TCCCCAAGA GGAAGGAAC | 4800 |
| CTTCTGTCGT CATATCTGGA AAGTTTAAAA TCCTAAATGT CATAAAGATT CCCAGACCTA | 4860 |
| GAATAGCCCA GACAAATCCT TGAGAAATAA TGGAAACAAT CATATTTTAT TTAATCCTTT | 4920 |

1070

| | |
|--|------|
| CTATATTCAT CTTTTTAAAA AATGGGAAGA GTCTCCTCCT CCCTACCTTA TTTATTCGAT | 4980 |
| GACTTGTCCT GCTTCTTTGA GAACAGACTC AGGAATAGTA ATACCTAGTT CTTGTGCTAT | 5040 |
| TTTTTTTATTG ATGACTGACT TACCAGTTGA AAAGACATTG ACTGGGGTAT CGGCTGGTTT | 5100 |
| TGCACCTTTC AAGACTTGCA CAATCATTTT ACCTGTTGCC ACACCAAGGT CATGTTGGTC | 5160 |
| AATTACAAC TATGCCAAAC CACCTACTTC TACCATAGCT GTCGCACTGG GATAAATTGG | 5220 |
| TTTCTTAGAA CTTTGATTGC TAGAGACAAC CGTTGGAAAT CCTGATGCAA TGGTGTATC | 5280 |
| AATTGGAACC CAAATAGCAT CTACCTTGCT AGTCATAACA GTGACAGTTG AGGCAATTTC | 5340 |
| ATTTGTTGAA GGAAGTCAA ATGTTTCCAC TGTCAGACCT GCCTTTTCAG CATAAGCCTT | 5400 |
| AAATTC | 5406 |

(2) INFORMATION FOR SEQ ID NO: 167:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 9711 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 167:

| | |
|--|-----|
| CAGCTTGCTC TTACTATTAT AGCAGATGTT ATAGCTGGAA TTATCTTGTA TTTCTGCTGC | 60 |
| AAATGGCTAG ATGGTAAGAA GTAGACCGAA TGAAGTACCT ATAAACACCC GTTAAATCGC | 120 |
| TAAGATACGT CAAAAAAGCC CTTAACTATG GCACTAGTTA GGGGCTTTGG TGTCTAATG | 180 |
| AACCTTATAC ACTAACTACA TTCTAGCATA TAAGCCCAGA TATTCAAGA GTTTTATTTA | 240 |
| TTGTTTAAAG TTCTGAAAGG TCTATAATGA AGTTAGCCAT CTAGTATCAA AAAACCGACT | 300 |
| AGCTCTTATG AACTAGTCGA TTTCTCATCA ATGCGCCAAC ATTTCTTGGG CGATTCTCTG | 360 |
| GCCAGATAGG TTATCTGGGT AGTAGGTTGG CCAAGTTGTC ATTTCTTCAA AGAGGGCTTC | 420 |
| TTGGCTTGTG CCTCCAAAGA AGATATGGAA ATGTTCTGCC TTAAGTGGG CAACATGTG | 480 |
| GTCACATAAC TGAACATACT TGAATTGTCC AGCGTCAGCA TCTGTGGCTT CAAAGAGGAA | 540 |
| ACGCACGCCA CGATTGCCTT TCTTGTAAGT CAAAATTTTC TTACCGACAT ACTTGTAAAGT | 600 |
| GTATTTCTTG CTTTGTCAC CTTGAACAAA TTCCATAGTA TTATCAGTAA TGTTAATCTT | 660 |
| AGTCACATCT GTATGATAGC CTTTGTGATA GTAAGCCTTG TACTCAGCCT GGGTCATCTT | 720 |
| ACCAGTCAAC TTAGCCTTGT AGTCAAAGAC TTGGTCAAAC GTGCCGTCTT CAAGGAAAGG | 780 |
| ATAAAGTGAT TGCCAGTTAC CTGCATAGTC ACTCAAGGTG CGGTCCTTGA CAGCTGCATC | 840 |
| CTCGAAGTAA CCATTTTGA CTGTCTGGT ATCCTCTGCC TTTTCAGGTT CAATTGCTGG | 900 |

1071

| | |
|--|------|
| GCCTTCTTGG TCTGTTGTTT GTTTCAAAGC CTTGAGGTTT TTCTCCATCA CGGAAATGTA | 960 |
| GTTTTCTCCA GCCTTGGTGT CCTCTTCTGT CAGACTTTCT AAAGGATTGA GGACATCAGT | 1020 |
| TTTGACACCT GCTTCTTTTG AAAGTGTGTT AGCAAGGGCT TGTGAGGCAT TTCTTCAAAA | 1080 |
| TAGATATAGG CGATTTTATT TTTCTTGACA TACTCTGTCA ATTCTGCCAA GCGAGCAGCT | 1140 |
| GATGGCTCTG CATCTGGAGA AAGTCCTGAG ATTGCGACTT GTTTGAGTCC ATAGTCCAAG | 1200 |
| GCAAGATAGT TAAAGGCTGC GTGTTGAGTC ACAAAGCTCT TTTGTTTTGC TTGAGACAAA | 1260 |
| CCTTCTGCGT AAGCCTTATC CAAGGCTTGC AATTTTTCGA TATAGGCAGC TGCATTCTTC | 1320 |
| TCAAAGGTCT CTTTTTTATC AGGATAATCT GCTGACAAGC TGTCGCGGAT GTGCTCTACT | 1380 |
| AGTTTAATGG CACGAACTGG TGATAACCAA ACATGGGGGT CAAACTCATG GTGATGACCT | 1440 |
| TCTTCTCCAT GGTCAATGGT TCCCTCTTCT TCCTCGCCAC CTGGCAAGAG CAACATATCG | 1500 |
| CCTGTGCGCT TGATGGTTTT CACTTTTTTC TTATCCAAGG TATCTAGCAA TTTAGGTACC | 1560 |
| CATGTTTCCA TGTTTTCATT TTCATAAACG AAGGTATCTG CATCTTGGAT TTTGGCAACT | 1620 |
| GCCTTGGCAG ATGGTTCGTA TTCATGAGGT TCTGTCCCAG CACCGATTAG GAGTCTTACA | 1680 |
| TTAGCCGTAT CTCCTGCGAC TTGCTTGGA AATTCATAGA CAGGGTAAAA GGTGTGCACG | 1740 |
| ATATTGAGTT TACCATCTGC CTGTTTTGA TTGGAACAAG CCACTAAAA CAAGGCACAT | 1800 |
| AGACTGGCTA GTAATAAGCT AATTTTTTTC ACGTTCGTCT CCTATTGAT AAAACGTCTT | 1860 |
| ACTAACTGA TTAGTATAAA GACAGTTACA AAAATAATGG TAATACTTGC ACTTGCAAGT | 1920 |
| GTTTCTGCAT AGTAGGAAAT GTAAAGTCCT GCTACCATT CCAAAAAGCC AATCGCACTG | 1980 |
| GCAAGCAGCA TAACCGATT AAAGTTTTTC CCCAGACGCA GGGCAATACT AGCTGGCAAG | 2040 |
| ACCATAATGG TCGATACCAG AAGAGCTCCT GCTGCAGGAA TCATAAGGGC AATAGCCACC | 2100 |
| CCTGTCACCA TGTTAAAAAG AATGGACATG GTACGAACTG GCAAGCCATC CACAAAGGCC | 2160 |
| GTATCTTCGT CAAAAGTTAA GATATACATA GGACGAAGAA AGAGAAAGGT CAAAATCAAA | 2220 |
| ACAACCGCCG CAATGACAAA GAGGGAAATG ACCTGTTCTT CACTGATAGT CACGATCGAA | 2280 |
| CCAAAGAGAT ATTGGTCCAA ACTCATTGAA CTCGAGCTTT TACCCCTGCT CATGACAATC | 2340 |
| AGAGAAACAG CCAGACCTGT TGACATGAGG ATAGCTGTCC CGATTTCCAT AAAGCTCTTG | 2400 |
| TAAACCGTAC GGAGATACTC CAGAAAGACC GCCGCAATCA AGACAATGGC AATAGTAGAA | 2460 |
| ACAGTTGGAG AAATCCCCAA AACCAGACCA AAGGCTACAC CTGAAAGTGA GACGTGGCTA | 2520 |
| AGGGTATCAC TCATCAAACCT CTGACGACGC AAGATGAGGA AGGTTCCCAA TACCGGTGAG | 2580 |
| AAAAGACTCA TAGCAATAAC CGCCAAAAAG GCGCGTTGTA TAAAGTCGTA AGATAATAAA | 2640 |

1072

| | |
|--|------|
| CTAAGCATGG CCCACCTCCT GGCCATTCTC ATGAACATTG AAACAACGCC ATGGCGAGTC | 2700 |
| TTGGTTACGG ACTAGATGAA TATTGCGATC CGCATAATCC TTAACCTCTT CAGGGTCATG | 2760 |
| GGTAATCATC AAAACAGCCT TGCCATGATG ATGGGCGCTG TGGTGCAATG GTTCGTAAAA | 2820 |
| TTCATTTTTC CTTCCTGCAT CCATCCCCGT TGTCGGCTCG TCTAGGATAA ACACATCAGG | 2880 |
| GTCAGAAGCA AACATACGCG CAATTACCGC TCGCTGCTTT TGTCCCCCAG ATAGAGACCC | 2940 |
| CAAGCGTTTG TCTCGATGTT CCCACATGCC AACTGAGTCC AGACTAGCCT TGATATGCTC | 3000 |
| CTCATCATGA GCATTCAAAC GACGGAACCA GCCTTTTCTC GGATAGCGAC CCGACTTGAC | 3060 |
| AAATTCATAG ACCGTAATG GAAAACCAGC ATTAAACTG GCAATTTGTT GAGGAAGATA | 3120 |
| GGCTATTCTC AATTTCTTAC CTGCGTATT TGCTTTTGAA ATAGCCACCT TTCCAATGCG | 3180 |
| TGGTTGCAGA ATTCCAAGAC TAGCCTTGAT GAGCGTCGTC TTAGCCGCTC CATTTTCCCC | 3240 |
| AGTCAAGGTA ACAAATTCCT CACTATCAAC ACAATAATTG ATATGTTCAA GAACAGGCTC | 3300 |
| CTTATCATAA TAGAAGGACA AATCCTCTAC CGTAATATAT CTCATTATTT GATTCTCTCT | 3360 |
| ACTAAAGCAG TCAAAAACCG CTGAATCACT TTTTGTTTAT TTGGAGTAAA CTGAGTCGCC | 3420 |
| ACTTGTTCAT AGGTAAAAAG TGTATGCTCA TGGTGATGGT GGTGCTCCTC AGCGATTGGA | 3480 |
| CGAGCCAAAGT CAGTCAACTG ATAAAAAATC ACACGCGCAT CTTTAGAATC TTTAGATGTT | 3540 |
| TCCAACATCC CTTCTTGAC CAAAGACTTA ATGGCCTTGG TAACTGCCGC CTGACTGACA | 3600 |
| TTGAGACGAC GGGCCAATTC TGAATTTGTT AAAGATTCTT CTGACAAGAG CATAAGGATA | 3660 |
| TGCTCCTGAG TATTGGTCAG GGCCACCTCG CTAGTGCAAT GACCTATTAG GATTTCATGC | 3720 |
| TGATTTTCCG CCTGCAAAAT CACCTCATTC AAAAAAGCAT TGATATCCTT TGCTAGCTGT | 3780 |
| CTCATATCTG ACTCCTTTCC TTTTAGACTT CTCTTTTTTA AGAGAAAAAT ACTATTCTTT | 3840 |
| GACATTTTGT TTACCAGTTA ATTATATCAC AAGCAAAAAA AGAGTCAAGA AAAAACGTGA | 3900 |
| AAACTAGTTT CATTCCTGAA CTCTTCTATA TTATATTATC TATTGAAATT CTTTGACATC | 3960 |
| TCCATCATAA GTCGCCCCAT CTTTGCTGAA AAAGCGCTCA TTCAGATGGT AAGTCGGAGC | 4020 |
| TGGTGTGGGA TTGGATAGGA AAGGATCAAC TGCCTTGTC AAGCCAACC AACCCAACCA | 4080 |
| ACCAAGGTGA ATGGTGTCTT TCATAAAGAA AGGCTCCCCG CCGTCTTAG AAAAATCTGC | 4140 |
| TATATTGGTA AAACCTTGAC TTTCTAACTG GTAGCGAATC TTCTGCACCG TTTGTTGGTA | 4200 |
| CATATCCTCT CGTAGACCAG CATAGTTCAT CCATTTTTC TTAACAGGTG GAATGATAAA | 4260 |
| AATCGGGTTT ACCTTAGATT TAGAAAACCTG TGTTAAAACC AACTGCAAGT CATTATACTC | 4320 |
| TGGCGACTTG AGATAGGTAA AGCTTTTCTG AGAATCCTTT AATTCTTCA AATCCTTCTT | 4380 |
| GATCTGCTCA TTATAGAAAT AATTTTCCAT TCCCATCTCA TTATTGGAAG TATTTTTC | 4440 |